

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 9.51698 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-55
Perfect score: 124
Sequence: 1 EOASVKYVILDMYRALTLMTSTA 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44.5	35.9	798	4	US-09-861-451A-12
2	44	35.5	278	1	US-07-921-807B-10
3	44	35.5	278	1	US-08-441-944A-10
4	43	34.7	428	4	US-09-328-352-5343
5	42	33.9	301	3	US-09-134-001C-4305
6	42	33.9	332	4	US-09-710-279-2350
7	41	33.1	248	4	US-09-248-796A-15839
8	41	33.1	425	1	US-08-700-749A-5
9	41	33.1	425	3	US-09-020-684-5
10	41	33.1	425	3	US-09-020-467-5
11	41	33.1	425	3	US-09-020-685-5
12	41	33.1	425	3	US-09-020-683-5
13	41	33.1	619	4	US-09-270-767-34769
14	41	33.1	619	4	US-09-270-767-49986
15	40.5	32.7	740	4	US-09-489-039A-11157
16	40	32.3	65	4	US-09-328-352-7222
17	40	32.3	161	2	US-08-403-852D-30
18	40	32.3	161	3	US-08-510-646B-31
19	40	32.3	161	3	US-09-231-818-30
20	40	32.3	161	4	US-09-635-359B-30
21	40	32.3	177	4	US-09-614-912-136
22	40	32.3	305	4	US-09-328-352-6794
23	40	32.3	527	4	US-09-538-092-238
24	40	32.3	531	4	US-09-252-991A-27660
25	39.5	31.9	65	4	US-09-248-796A-27954
26	39.5	31.9	345	4	US-09-270-767-60654
27	39.5	31.9	404	4	US-09-270-767-45164

28	39.5	31.9	998	4	US-09-540-236-3801	Sequence 3801, Ap
29	39.5	31.9	1724	2	US-08-477-451-15	Sequence 15, Appl
30	39	31.5	112	4	US-09-248-796A-26280	Sequence 26280, A
31	39	31.5	178	4	US-09-543-681A-7128	Sequence 7128, Ap
32	39	31.5	236	4	US-09-107-532A-6356	Sequence 6356, Ap
33	39	31.5	305	4	US-09-340-236-3362	Sequence 3362, Ap
34	39	31.5	317	4	US-09-248-796A-20568	Sequence 20568, A
35	39	31.5	370	2	US-08-837-593-7	Sequence 7, Appl
36	39	31.5	371	2	US-08-837-593-6	Sequence 6, Appl
37	39	31.5	375	2	US-08-837-593-5	Sequence 5, Appl
38	39	31.5	375	4	US-09-623-034-2	Sequence 2, Appl
39	39	31.5	387	2	US-08-837-593-4	Sequence 4, Appl
40	39	31.5	393	2	US-08-837-593-3	Sequence 3, Appl
41	39	31.5	431	4	US-09-270-767-42950	Sequence 42950, A
42	39	31.5	434	4	US-09-538-092-209	Sequence 209, App
43	39	31.5	450	4	US-09-721-870-16	Sequence 16, Appl
44	39	31.5	551	4	US-09-710-279-1668	Sequence 1668, Ap
45	39	31.5	560	3	US-09-134-001C-3153	Sequence 3153, Ap

ALIGNMENTS

RESULT 1
US-09-861-451A-12
; Sequence 12, Application US/09861451A
; Patent No. 6759516
; GENERAL INFORMATION:
; APPLICANT: Commonwealth Scientific & Industrial Research Orga
; TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences
; FILE REFERENCE: PF34033/01
; CURRENT APPLICATION NUMBER: US/09/861,451A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PF7273
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 798
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced protein
; OTHER INFORMATION: sequence from clone PAD653
US-09-861-451A-12

Query Match 35.9%; Score 44.5; DB 4; Length 798;
Best Local Similarity 39.3%; Pred.No.76;
Matches 11; Conservative 7; Mismatches 7; Indels 3; Gaps 1;

QY 1 EOASVKYVIL--DMYRALTLMTSTA 25
DB 605 EATKFNILVAESAYQIVSLFNNSNA 632

RESULT 2
US-07-921-807B-10
; Sequence 10, Application US/07921807B
; Patent No. 5474314
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

; MOLECULE TYPE: protein
US-08-441-944A-10

Query Match      35.5%; Score 44; DB 1; Length 278;
Best Local Similarity 60.0%; Pred.No. 27;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;


QY 11 DMYRALLTLMNTSTA 25
| | | | | : : : |
Db 117 DQLRALLTLLSSDTA 131

RESULT 4
US-09-328-352-5543
; Sequence 5543, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5543
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5543

Query Match      34.7%; Score 43; DB 4; Length 428;
Best Local Similarity 42.9%; Pred.No. 65;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;


QY 5 VKYVILDMYRALLTLMNTSTA 25
| | | | | : : : |
Db 89 VPYITLDFHCLLSAMNTALA 109

RESULT 5
US-09-134-001C-4305
; Sequence 4305, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4305
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4305

Query Match      33.9%; Score 42; DB 3; Length 301;
Best Local Similarity 30.4%; Pred.No. 63;
Matches 7; Conservative 8; Mismatches 8; Indels 0; Gaps 0;


QY 2 QASVKYVILDMYRALLTLMNTST 24
| | | | | : : : |
Db 273 BEAKHLIFDALDAMEVLNTGT 295

RESULT 6
US-09-710-279-2350
; Sequence 2350, Application US/09710279

```

```
, STATE: California
, COUNTRY: USA
, ZIP: 94080
, COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,749A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/002382
FILING DATE: 17-aug-1995
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P0960R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-3216
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-700-749A-5

Query Match      33.1%; Score 41; DB 1; Length 425;
Best Local Similarity 34.8%; Pred.No.1.4e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
```

QY 1 EQASVKVYLDMYRALTLNMTS 23
 | | : | : | : | : | :
Db 46 EQLSQQTIDKLKSLLVNST 68

```
RESULT 9
US-09-020-684-5
Sequence 5, Application US/09020684
Patent No. 6004553
GENERAL INFORMATION:
APPLICANT: GOEDEL, DAVID V.
APPlicant: ROTHE, MIKE
TITLE OF INVENTION: TRAF INHIBITORS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,684
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/002382
FILING DATE: 17-aug-1995
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P0960RID2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-3216
TELEFAX: 650/952-9881
```

```
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-020-684-5

Query Match      33.1%; Score 41; DB 3; Length 425;
Best Local Similarity 34.8%; Pred. No. 1.4e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMYRALLTLNMTS 23
Db 46 EQLSLOQTIDKLKSQLLLVNST 68

RESULT 10
US-09-020-467-5
; Sequence 5, Application US/09020467
; Patent No. 6060303
; GENERAL INFORMATION:
; APPLICANT: GOEDDEL, DAVID V.
; APPLICANT: ROTHE, MIKE
; TITLE OF INVENTION: TRAF INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,467
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002382
; FILING DATE: 17-aug-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P0960R1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-020-685-5

Query Match      33.1%; Score 41; DB 3; Length 425;
Best Local Similarity 34.8%; Pred. No. 1.4e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMYRALLTLNMTS 23
Db 46 EQLSLOQTIDKLKSQLLLVNST 68

RESULT 12
US-09-020-683-5
; Sequence 5, Application US/09020683
; Patent No. 6294348
; GENERAL INFORMATION:
; APPLICANT: GOEDDEL, DAVID V.
; APPLICANT: ROTHE, MIKE
; TITLE OF INVENTION: TRAF INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,683
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002382
; FILING DATE: 17-aug-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, DAVID V.
```


ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P0960R1D3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-3216
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-020-683-5

Query Match 33.1%; Score 41; DB 3; Length 425;
Best Local Similarity 34.8%; Pred. No. 1.4e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMYRALTLMTSTAT 23
DB 46 EQLSLQQTIIIDKLSQLLVNST 68

RESULT 13
US-09-270-767-34769
Sequence 34769, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 34769
LENGTH: 619
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34769

Query Match 33.1%; Score 41; DB 4; Length 619;
Best Local Similarity 34.5%; Pred. No. 2.1e+02;
Matches 10; Conservative 6; Mismatches 9; Indels 4; Gaps 1;

QY 2 QASVKYVILDMYRAL---LTLMTSTAT 26
DB 418 QGAYDYTYIEIYRYLYIFHFTCVTSFST 446

RESULT 14
US-09-270-767-49986
Sequence 49986, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 49986
LENGTH: 619
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49986

Query Match 33.1%; Score 41; DB 4; Length 619;

Best Local Similarity 34.5%; Pred. No. 2.1e+02;
Matches 10; Conservative 6; Mismatches 9; Indels 4; Gaps 1;

QY 2 QASVKYVILDMYRAL---LTLMTSTAT 26
DB 418 QGAYDYTYIEIYRYLYIFHFTCVTSFST 446

RESULT 15
US-09-489-039A-11157
Sequence 11157, Application US/09489039A
Patent No. 6610836

GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 11157
LENGTH: 740
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11157

Query Match 32.7%; Score 40.5; DB 4; Length 740;
Best Local Similarity 32.1%; Pred. No. 3.1e+02;
Matches 9; Conservative 7; Mismatches 3; Indels 9; Gaps 1;

QY 3 ASVKYVILDMYRA-----LTLTLMN 21
DB 156 SSVAYIVIDEFHAFISGSRGVQLLSLN 183

Search completed: November 10, 2004, 14:55:40
Job time : 10.567 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 30.0717 Seconds
(without alignments)

305.399 Million cell updates/sec

Title: US-10-092-750-55

Perfect score: 124

Sequence: 1 EQASVKVILDMYRALTLMTSTAT 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35322586 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/FCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	26	14	Sequence 55, Appl
2	48.5	39.1	398	15	Sequence 185720
3	48	38.7	357	16	Sequence 136288
4	47.5	38.3	122	15	Sequence 254656
5	47	37.9	95	11	Sequence 8064, Ap
6	45	36.3	78	17	Sequence 311380
7	45	36.3	1205	15	Sequence 2967, Ap
8	44.5	35.9	798	9	Sequence 12, Appl
9	44	35.5	91	15	Sequence 271940
10	44	35.5	109	15	Sequence 270166
11	44	35.5	737	15	Sequence 256196
12	43.5	35.1	307	9	Sequence 6331, Ap
13	43	34.7	103	16	Sequence 38851, A

14	43	34.7	420	15	US-10-282-122A-44699	Sequence 44699, A
15	43	34.7	477	15	US-10-424-599-175572	Sequence 175572, A
16	43	34.7	858	15	US-10-425-114-72451	Sequence 72451, A
17	43	34.7	872	15	US-10-424-599-219651	Sequence 219651, A
18	42.5	34.3	237	16	US-10-741-131-35	Sequence 35, Appl
19	42.5	34.3	237	16	US-10-742-350-35	Sequence 35, Appl
20	42	33.9	43	15	US-10-424-599-186763	Sequence 186763, A
21	42	33.9	43	15	US-10-424-599-212841	Sequence 212841, A
22	42	33.9	108	17	US-10-425-115-361202	Sequence 361202, A
23	42	33.9	138	17	US-10-425-115-284221	Sequence 284221, A
24	42	33.9	206	17	US-10-425-115-284217	Sequence 128472, A
25	42	33.9	352	16	US-10-437-963-125472	Sequence 125472, A
26	42	33.9	359	17	US-10-425-115-284222	Sequence 284222, A
27	42	33.9	380	15	US-10-425-114-48492	Sequence 48492, A
28	42	33.9	380	15	US-10-425-114-66515	Sequence 66515, A
29	42	33.9	380	15	US-10-425-114-70564	Sequence 70564, A
30	42	33.9	1071	16	US-10-437-963-176644	Sequence 176644, A
31	41.5	33.5	129	17	US-10-425-115-245548	Sequence 245548, A
32	41.5	33.5	231	17	US-10-425-115-284441	Sequence 328441, A
33	41	33.1	65	15	US-10-424-599-276237	Sequence 276237, A
34	41	33.1	119	15	US-10-424-599-197293	Sequence 197293, A
35	41	33.1	119	17	US-10-788-792-189	Sequence 189, App
36	41	33.1	154	17	US-10-425-115-276085	Sequence 276085, A
37	41	33.1	160	15	US-10-425-114-51400	Sequence 51400, A
38	41	33.1	227	15	US-10-424-599-284783	Sequence 284783, A
39	41	33.1	281	16	US-10-767-701-41939	Sequence 41939, A
40	41	33.1	287	14	US-10-263-568-9	Sequence 9, Appl
41	41	33.1	347	15	US-10-425-114-73116	Sequence 73116, A
42	41	33.1	362	16	US-10-437-963-115286	Sequence 115286, A
43	41	33.1	381	14	US-10-369-493-19997	Sequence 19997, A
44	41	33.1	409	14	US-10-369-493-10754	Sequence 10754, A
45	41	33.1	425	14	US-10-247-671-148	Sequence 148, App

ALIGNMENTS

RESULT 1
US-10-092-750-55
; Sequence 55, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-55

Query Match 100.0%; Score 124; DB 14; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EQASVKVILDMYRALTLMTSTAT 26
Db 1 EQASVKVILDMYRALTLMTSTAT 26

RESULT 2
US-10-424-599-185720
; Sequence 185720, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

```

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285694
; SEQ ID NO 185720
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(398)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_138719C.1.dep
; US-10-424-599-185720

```

```
Query Match      39.1%; Score 48.5; DB 15; Length 398;
Best Local Similarity 46.2%; Pred. NO. 24;
Matches 12; Conservative 5; Mismatches 8; Indels 1; Caps 1;
```

```

RESULT 3
US-10-437-963-136288
US-10-437-963-136288, Application US/10437963
; Sequence 136288, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 136288
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37882C.1.pbp
US-10-437-963-136288

```

Query Match 38.7%; Score 48; DB 16; Length 357;
Best Local Similarity 36.0%; Pred. No. 26;
Matches 9; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

RESULT 4
US-10-424-599-254556
; Sequence 254556, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J

```

; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCES: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254656
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(122)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3947_71978C.1.pep
; US-10-424-599-254656

```

Query Match 38.3%; Score 47.5; DB 15; Length 122;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

```

RESULT 5
US-09-864-408A-8064
; Sequence 8064, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Enc
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8064
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-09-864-408A-8064

```

```
Query Match      37.9%; Score 47; DB 11; Length 95;
Best Local Similarity 34.8%; Pred. No. 8.2;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
```

RESULT 6
US-10-425-115-311380
; Sequence 311380, Application US/10425115
; Publication No. US20040314272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 311380
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_47035C.1.pep
US-10-425-115-311380

Query Match      36.3%; Score 45; DB 17; Length 78;
Best Local Similarity 46.2%; Pred. No. 14;
Matches 12; Conservative 2; Mismatches 4; Indels 8; Gaps 1;

Qy 9 ILDMYRAL-----LTLMNTSTAT 26
Db 9 VLDYRSRLRQVTDVQLTCMRTFTAT 34

RESULT 7
US-10-108-260A-2967
; Sequence 2967, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2967
; LENGTH: 1205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2967

Query Match      36.3%; Score 45; DB 15; Length 1205;
Best Local Similarity 38.5%; Pred. No. 3,3e+02;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 EQASVKYVILDMYRALTLTMTSTAT 26
Db 282 EKSLKTAVISGNQLHLKETDTAT 307

RESULT 8
US-09-861-451A-12
; Sequence 12, Application US/09861451A
; Patent No. US20020068289A1
; GENERAL INFORMATION:
; APPLICANT: Commonwealth Scientific & Industrial Research Orga
; TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences
; FILE REFERENCE: FF34033/01
; CURRENT APPLICATION NUMBER: US/09/861,451A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PP7273
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 798
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced protein
; OTHER INFORMATION: sequence from clone PAD653
US-09-861-451A-12

```

```

Query Match      35.9%; Score 44.5; DB 9; Length 798;
Best Local Similarity 39.3%; Pred. No. 2.5e+02;
Matches 11; Conservative 7; Mismatches 7; Indels 3; Gaps 1;

Qy 1 EQASVKYVIL---DMYRALTLTMTSTA 25
Db 605 BEATKKFNILVAESAYQAIIVSLFNNSNA 632

RESULT 9
US-10-424-599-271940
; Sequence 271940, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 271940
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87580C.1.pep
US-10-424-599-271940

Query Match      35.5%; Score 44; DB 15; Length 91;
Best Local Similarity 45.0%; Pred. No. 24;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 6 KVVILDMYRALTLTMTSTA 25
Db 67 KVVILDTTCLMILLNKYTS 86

RESULT 10
US-10-424-599-270166
; Sequence 270166, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 270166
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_8597C.1.pep
US-10-424-599-270166

Query Match      35.5%; Score 44; DB 15; Length 109;
Best Local Similarity 47.6%; Pred. No. 30;
Matches 10; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

Qy 1 EQASVKYVILDMYRALTLTMTN 21
Db 17 EQSSLGF--LDLQALLVFFN 35

```

RESULT 11

US-10-424-599-256196
 ; Sequence 256196, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 256196
 ; LENGTH: 737
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT MRT3847_73366C.1.pep
 US-10-424-599-256196

Query Match 35.5%; Score 44; DB 15; Length 737;
 Best Local Similarity 66.7%; Pred. No. 2.7e-02;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 10 LMYRALLTMTST 24
 ||| |||:|:|:|
 Db 146 LMELELLTMTLT 160

RESULT 12

US-09-738-626-6331
 ; Sequence 6331, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIALI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 6331
 ; LENGTH: 307
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-6331

Query Match 35.1%; Score 43.5; DB 9; Length 307;
 Best Local Similarity 44.0%; Pred. No. 1.2e+02;
 Matches 11; Conservative 5; Mismatches 4; Indels 5; Gaps 1;

QY 2 QASVKYVILDMYRALLTMTSTAT 26
 |||: :|||:|:|:|
 Db 255 QASPLMLDMYKA-----GTAPAT 274

RESULT 13

US-10-767-701-38851
 ; Sequence 38851, Application US/10767701
 ; Publication No. US20040172684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 ; FILE REFERENCE: 38-21(53535)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-29
 ; NUMBER OF SEQ ID NOS: 63128
 ; SEQ ID NO 38851
 ; LENGTH: 103
 ; TYPE: PRT
 ; ORGANISM: Sorghum bicolor
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C76631_1.pep
 US-10-767-701-38851

Query Match 34.7%; Score 43; DB 16; Length 103;
 Best Local Similarity 36.4%; Pred. No. 41;
 Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 3 ASVKYVILDMYRALLTMTST 24
 :|:|:|:|:|:|
 Db 16 SDLKLVVFDIYQFLDLVQNPT 37

RESULT 14

US-10-282-122A-44699
 ; Sequence 44699, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44599
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-44599

Query Match      34.7%; Score 43; DB 15; Length 420;
Best Local Similarity 42.9%; Pred. No. 2.1e+02;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy      5 VKYVILDMYRALTLMTSTA 25
Db      81 VPIITLDFIHCLLSAMNIALA 101

RESULT 15
US-10-424-599-175572
; Sequence 175572, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175572
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12955C.1.pep
US-10-424-599-175572

Query Match      34.7%; Score 43; DB 15; Length 477;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Qy      1 EQASVKYVILDMYRALTLMTSTAT 26
Db      182 EQRSAPYSMDRSRAVYTOQNNVIGT 207

Search completed: November 11, 2004, 02:43:11
Job time : 31.1217 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 6.37736 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-55

Perfect score: 124

Sequence: 1 EQASVKYVLDYMYRALLTLMNTSTAT 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	51	41.1	345	2	F86368	hypothetical prote
2	50	40.3	389	2	E90431	sulfolipid biosynt
3	47	37.9	1137	2	A86335	T20H2.9 protein -
4	44	35.5	295	2	A99489	arabinose ABC tran
5	44	35.5	306	2	S09882	hypothetical prote
6	44	35.5	1242	2	T39453	probable mrna stab
7	43	34.7	172	2	S69749	hypothetical prote
8	43	34.7	195	2	H70171	hypothetical prote
9	43	34.7	359	2	T48388	hypothetical prote
10	43	34.7	426	1	B69656	multidrug-efflux t
11	43	34.7	483	2	AF3373	glutathione-disulf
12	43	34.7	503	2	D75104	transmembrane tran
13	43	34.7	729	2	T52187	probable transpos
14	42.5	34.3	237	2	A55218	sfp homolog gsp -
15	42.5	34.3	474	2	F75133	hypothetical prote
16	42	33.9	282	2	G95130	transcription regu
17	42	33.9	452	2	C65140	conserved hypothet
18	42	33.9	483	2	S45347	probable G protein
19	42	33.9	521	2	AB0269	anthranilate synth
20	42	33.9	539	2	A70899	probable PPE prote
21	42	33.9	918	2	S45872	hypothetical prote
22	42	33.9	1309	2	S62457	hypothetical prote
23	42	33.9	1318	2	T38568	hypothetical prote
24	42	33.9	3973	2	D71612	hypothetical prote
25	41.5	33.5	160	2	AD2490	hypothetical prote
26	41.5	33.5	236	2	F75373	probable DNA-(apur
27	41	33.1	231	2	T48386	hypothetical prote
28	41	33.1	391	2	T32156	hypothetical prote
29	41	33.1	445	2	H71176	probable signal re

sulfolipid biosynt
hypothetical prote
Tik7.22 protein -
hypothetical prote
DNA ligase (ATP) (
probable transport
hypothetical DNA-b
hypothetical prote
protein F30B17.4 (
multidrug resistan
cytochrome-c oxida
lectin precursor -
arcelin 5a precurs
arcelin 5b precurs
conserved hypothet

ALIGNMENTS

RESULT 1

F86368

hypothetical protein F508.6 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: F86368

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
Xer, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F86368

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-345 <STO>

A;Cross-references: UNIPROT:Q9ZUE4; GB:AE005172; NID:g4056433; PIDN:AAC98006.1; GSPDB:G

C;Genetics:

A;Map position: 1

C;Superfamily: myrosinase-associated protein MYAP

Query Match 41.1%; Score 51; DB 2; Length 345;
Best Local Similarity 47.6%; Pred. No. 2.1;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 6 KYVLDYMYRALLTLMNTSTAT 26

Db 271 KFYVVDYMYRLLNINPRAS 291

RESULT 2

E90431

sulfolipid biosynthesis protein (sgdB) [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C;Accession: E90431

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: E90431

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-389 <KUR>

A;Cross-references: UNIPROT:Q97VN5; GB:AE006641; NID:gl3815887; PIDN:AAK42708.1; GSPDB

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 : Search time 34.634 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-55

Sequence: 1 EQASVKYVILDMYRALLTLMNTSTAT 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	53	42.7	317	2	Q8RGL6		Q8rgl6 fusobacteri
2	52	41.9	274	1	VGLL_MCMVK		P52513 murine cyto
3	52	41.9	274	1	VGLL_MCMVS		P52514 murine cyto
4	51	41.1	345	2	Q9ZUE4		Q9zue4 arabidopsis
5	51	41.1	375	2	Q9LQC7		Q9lqc7 arabidopsis
6	51	41.1	428	2	O8RFT8		O8rft8 fusobacteri
7	50	40.3	389	2	O9TVN5		O9tvn5 sulfolobus
8	48	38.7	237	1	THX_BPPHC		O9rx92 bacterioph
9	48	38.7	357	2	Q7F007		Q7f007 oryza sativ
10	48	38.7	357	2	Q9AUJ9		Q9auj9 oryza sativ
11	47.5	38.3	149	2	Q9K2L9		Q9k2l9 erwinia tra
12	47	37.9	637	2	O6D999		Q6d999 erwinia car
13	47	37.9	680	2	Q7XPT9		Q7xpt9 oryza sativ
14	46	37.1	262	2	O8SKT7		O8skt7 dictyosteli
15	46	37.1	641	2	O8PST9		O8pst9 methanosarc
16	46	37.1	694	2	Q8IK91		Q8ik91 plasmodium
17	46	37.1	1065	2	Q7TDB4		Q7tdb4 cryphonectr
18	45	36.3	158	2	Q8IIN4		Q8iin4 plasmodium
19	45	36.3	270	2	Q7P4T7		Q7p4t7 fusobacteri
20	45	36.3	299	2	Q38246		Q38246 lactococcus
21	45	36.3	299	2	Q38292		Q38292 lactococcus
22	45	36.3	661	2	Q6CMY7		Q6cmv7 kluyveromyc
23	45	36.3	884	2	Q9MA60		Q9ma60 arabidopsis
24	45	36.3	6885	1	SNE2_HUMAN		Q8wxh0 homo sapien
25	44.5	35.9	606	2	Q7KVA9		Q7kva9 drosophila
26	44.5	35.9	606	2	A4F47526		Aaf47526 drosophila
27	44.5	35.9	709	2	Q7KVB1		Q7kvb1 drosophila
28	44.5	35.9	709	2	A4F47524		Aaf47524 drosophila
29	44.5	35.9	743	2	Q9W0C5		Q9w0c5 drosophila
30	44.5	35.9	743	2	A4F47525		Aaf47525 drosophila
31	44.5	35.9	1398	2	Q9XZ32		Q9xz32 drosophila

ALIGNMENTS

RESULT 1

ID Q8RGL6 PRELIMINARY; PRT; 317 AA.
AC Q8RGL6;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Transposase.
GN OrderedLocusNames=FN0275;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 25586;
RC MEDLINE=21886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Fusch G., Haselkorn R.,
RA Fontstein M., Kyripides N.C., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EXML; AE010539; AAL94481.1; -
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:Transposase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR009057; Homeodomain like.
DR InterPro; IPR002560; Transposase_12.
DR Pfam; PF01610; Transposase_12; 1.
DR Complete proteome.
KW SEQUENCE 317 AA; 37031 MW; 716DCAE60ABFC885 CRC64;
SQ

Query Match 42.7%; Score 53; DB 2; Length 317;
Best Local Similarity 42.1%; Pred. No. 3.5;
Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 4 SVKXVILDMYRALLTLMNT 22

DB 219 NVKYICMDMTYSISLVNS 237

RESULT 2

ID VGLL_MCMVK STANDARD; PRT; 274 AA.
AC VGLL_MCMVK;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Glycoprotein L precursor.
GN Name=GL; Synonyms=UL15;
OS Murine cytomegalovirus (strain KI81).

Q6C3P1 Yarrowia li
Q68667 human cytom
Q68668 human cytom
Q68669 human cytom
Q68670 human cytom
Q68671 human cytom
Q68672 human cytom
Q68673 human cytom
Q68674 human cytom
P16932 human cytom
Q68666 human cytom
Q714D9 human cytom
Q714P4 human cytom
Q7M6S9 human cytom

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Muromegalovirus.
 OX NCBI_TaxID=69156;
 RN [1]
 RP MEDLINE=95053910; PubMed=7964634;
 RX Xu J., Scalzo A.A., Lyons P.A., Farrell H.E., Rawlinson W.D.,
 RA Shellam G.R.;
 RT "Identification, sequencing and expression of the glycoprotein L gene
 of murine cytomegalovirus.";
 RL J. Gen. Virol. 75:3235-3240(1994).
 CC -!- FUNCTION: Associated with glycoprotein H (gH) to form a complex
 CC important for infection and cell fusion. This association is
 CC necessary for the correct processing and cell surface expression
 CC of gH (By similarity).
 CC -!- SIMILARITY: Belongs to the herpesviruses glycoprotein L family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L32962; AAA57343.1; -
 DR InterPro: IPR002689; Cytomegalo_gL.
 DR Pfam: PF01801; Cytomegalo_gL; 1.
 KW Envelope protein; Glycoprotein; Signal.
 FT SIGNAL 1 17 Potential.
 FT CHAIN 18 274 Glycoprotein L.
 FT CARBOHYD 48 48 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 64 64 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 116 116 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 170 170 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 274 274 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 274 AA; 31239 MW; 504D349FDBEC5C03 CRC64;
 Query Match 41.9%; Score 52; DB 1; Length 274;
 Best Local Similarity 50.0%; Pred. No. 4.4;
 Matches 13; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
 QY 1 EQASVKYVILDMYRALLTLMNTSTAT 26
 DB 109 EQSLHLNNTDQLRVLLTLMRTSRAS 134
 RESULT 3
 VGLL_MCMVS STANDARD; PRT; 274 AA.
 AC P52514;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Glycoprotein L precursor.
 GN Name=GL; Synonyms=UL15;
 OS Murine cytomegalovirus (strain Smith).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Muromegalovirus.
 OX NCBI_TaxID=10367;
 RN [1]
 RP MEDLINE=95053910; PubMed=7964634;
 RX Xu J., Scalzo A.A., Lyons P.A., Farrell H.E., Rawlinson W.D.,
 RA Shellam G.R.;
 RT "Identification, sequencing and expression of the glycoprotein L gene
 of murine cytomegalovirus.";
 RL J. Gen. Virol. 75:3235-3240(1994).
 CC -!- FUNCTION: Associated with glycoprotein H (gH) to form a complex
 CC important for infection and cell fusion. This association is
 CC necessary for the correct processing and cell surface expression
 CC of gH (By similarity).
 CC -!- SIMILARITY: Belongs to the herpesviruses glycoprotein L family.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L32963; AAA57344.1; -
 DR InterPro: IPR002689; Cytomegalo_gL.
 DR Pfam: PF01801; Cytomegalo_gL; 1.
 KW Envelope protein; Glycoprotein; Signal.
 FT SIGNAL 1 17 Potential.
 FT CHAIN 18 274 Glycoprotein L.
 FT CARBOHYD 48 48 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 64 64 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 116 116 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 170 170 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 274 274 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 274 AA; 31221 MW; 504D359DEB5D003 CRC64;
 Query Match 41.9%; Score 52; DB 1; Length 274;
 Best Local Similarity 50.0%; Pred. No. 4.4;
 Matches 13; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
 QY 1 EQASVKYVILDMYRALLTLMNTSTAT 26
 DB 109 EQSLHLNNTDQLRVLLTLMRTSRAS 134
 RESULT 4
 Q9ZUE4 PRELIMINARY; PRT; 345 AA.
 ID Q9ZUE4
 AC Q9ZUE4;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE F508.6 protein.
 GN Name=F508.6;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Liu S., Lenz C.,
 RA Li J., Kremenetskaia I., Luros J., Altafi H., Gonzalez A., Araujo R.,
 RA Buehler E., Conn L., Conway A.B., Dunn P., Hansen N., Huizar L.,
 RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Theologis;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC005990; AAC98006.1; -
 DR PIR; F86368; F86368.
 DR GO; GO:0003824; F: catalytic activity; IEA.
 DR InterPro; IPR001087; Lipase_GDSL.
 DR InterPro; IPR008285; Lipase_GDSL_AS.
 DR Pfam; PF00657; Lipase_GDSL; 1.
 DR PROSITE; PS01098; LIPASE_GDSL_SER; 1.
 SQ SEQUENCE 345 AA; 37472 MW; 98368ABBB60A9C1E CRC64;
 Query Match 41.1%; Score 51; DB 2; Length 345;
 Best Local Similarity 47.6%; Pred. No. 8.3;
 Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 6 KYVILDMYRALLTLMNTSTAT 26
 DB 271 KFVYVDMYNPLNLLNNPRAS 291

```

RESULT 5
Q9LQC7 PRELIMINARY; PRT; 375 AA.
AC Q9LQC7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F28C11.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC007945; AAF79588.1; -.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR InterPro: IPR001087; Lipase_GDSL.
DR InterPro: IPR008265; Lipase_GDSL_AS.
DR Pfam: PF00657; Lipase_GDSL; 1.
DR PROSITE: PS01098; LIPASE_GDSL_SER; 1.
SQ SEQUENCE 375 AA; 41143 MW; F1EFA4D13BB7243 CRC64;

Query Match 41.1%; Score 51; DB 2; Length 375;
Best Local Similarity 47.6%; Pred. No. 9.1;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 6 KVVILDMYRALTLMTSTAT 26
Db 301 KVVYVDMYNPLINPRAS 321

RESULT 6
Q8RFT8 PRELIMINARY; PRT; 428 AA.
AC Q8RFT8
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transposase..
GN OrderedLocusNames=FN0599;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;

```

```

OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Malunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyripides N.C., Overbeek R.;
RA "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010571; AAL94795.1; -.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004803; F:transposase activity; IEA.
DR GO: GO:0006310; P:DNA recombination; IEA.
DR InterPro: IPR009057; Homeodomain-like.
DR InterPro: IPR002560; Transposase_12.
DR Pfam: PF01610; Transposase_12; 1.
KW Complete proteome.
SQ SEQUENCE 428 AA; 49937 MW; BA99FA4348FAB3C3 CRC64;

Query Match 41.1%; Score 51; DB 2; Length 428;
Best Local Similarity 42.1%; Pred. No. 10;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 4 SVKYVILDMYRALTLMT 22
Db 219 NVKICMDMYTPISLVNS 237

RESULT 7
Q97VNS PRELIMINARY; PRT; 389 AA.
ID Q97VNS
AC Q97VNS;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sulfolipid biosynthesis protein (Sgdb).
GN Namesgdb: OrderedLocusNames=SSO2583;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
RL EMBL; AE006855; AAK42708.1; -.
DR PIR; E90431; E90431.
DR HSP; O48917; 1QRR.
KW Complete proteome.
SQ SEQUENCE 389 AA; 44686 MW; 1A63ADFF33A26C3 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 389;
Best Local Similarity 42.3%; Pred. No. 14;
Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMYRALTLMTSTAT 26
Db 99 EQRSAPYSMDVNHVYTLINLTST 124

```

RA	Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.
RA	White O., Salzberg S.L., Fraser C.M.;
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBSJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RA	Buell R.;
RL	Submitted (JUN-2003) to the EMBL/GenBank/DBSJ databases.
DR	EMBL; AC087412; AAP68380.1; -
DR	Gramene; Q7I007; -
DR	GO; GO:0003824; F:catalytic activity; IEA.
DR	InterPro; IPR001087; Lipase_GDSL.
DR	IntePro; IPR008285; Lipase_GDSL_AS.
DR	Pfam; PF00657; Lipase_GDSL; 1.
DR	PROSITE; PS01098; LIPASE_GDSL_SER; 1.
KW	Hypothetical protein.
SQ	SEQUENCE 357 AA; 37411 MW; 2F6C996795439C6C CRC64;
Query Match 38.7%; Score 48; DB 2; Length 357;	
Best Local Similarity 36.0%; Pred.No. 28;	
Matches 9; Conservative 6; Mismatches 10; Indels 0; Gaps 0;	
QY	1 EQASVKVILDMYRALLTLMNTSTA 25 : ::: :
Db	268 QHSDLKLVVDIYNPLLDIVNPTA 292 : : : :
RESULT 10	
Q9AUJ9	PRELIMINARY; PRT; 357 AA.
ID	AC Q9AUJ9;
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Putative proline-rich protein.
GN	Name=OSUNB007ZE24.7;
OS	Oryza sativa (rice).
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC	Ehrhartoidae; Oryzaceae; Oryza.
OX	NCBI_TaxID=4530;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Buell R.;
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
DR	EMBL; AC087851; AAK14416.1; -
DR	Gramene; Q9AUJ9; -
DR	GO; GO:0003824; F:catalytic activity; IEA.
DR	InterPro; IPR001087; Lipase_GDSL.
DR	IntePro; IPR008285; Lipase_GDSL_AS.
DR	Pfam; PF00657; Lipase_GDSL; 1.
DR	PROSITE; PS01098; LIPASE_GDSL_SER; 1.
SQ	SEQUENCE 357 AA; 37411 MW; 2F6C996795439C6C CRC64;
Query Match 38.7%; Score 48; DB 2; Length 357;	
Best Local Similarity 36.0%; Pred.No. 28;	
Matches 9; Conservative 6; Mismatches 10; Indels 0; Gaps 0;	
QY	1 EQASVKVILDMYRALLTLMNTSTA 25 : : : :
Db	268 QHSDLKLVVDIYNPLLDIVNPTA 292 : : : :
RESULT 11	
Q9KZL9	PRELIMINARY; PRT; 149 AA.
ID	AC Q9KZL9;

DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE Glyceralddehyde-3-phosphate dehydrogenase (Fragment).
 GN Name=gapDH;
 OS Erwinia tracheiphila.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Erwinia.
 OC NCBI_TaxID=65700;
 RN [1]_TaxID=65700;
 RP SEQUENCE FROM N.A.
 RC STRAIN=ET9815, and ET5845;
 RX MEDLINE=21027675; PubMed=11155980;
 RA Brown B.W., Davis R.M., Gouk C., van der Zwet T.;
 RT "Phylogenetic relationships of necrogenic Erwinia and Brenneria
 RT species as revealed by glyceraldehyde-3-phosphate dehydrogenase gene
 RT sequences.";
 RL Int. J. Syst. Evol. Microbiol. 50:2057-2068 (2000).
 CC -!- SIMILARITY: Belongs to the glyceraldehyde-3-phosphate
 CC dehydrogenase family.
 DR EMBL; AF165012; AAF31217.1; -;
 DR EMBL; AF165011; AAF31216.1; -;
 DR HSSP; P06977; 1DC5.
 DR GO; GO:0004365; P:glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.
 DR GO; GO:0064591; P:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR000173; GAP_dhhdhrogenase.
 DR Pfam; PF02800; Gp dh C; 1.
 DR Pfam; PF00044; Gp dh N; 1.
 DR PRINTS; PR00078; G3PDHDSGNASE.
 KW Oxidoreductase.
 FT NON_TER 1
 FT NON_TER 149
 SQ SEQUENCE 149 AA; 16376 MW; D89D17E8958E0122 CRC64;
 Query Match 38.3%; Score 47.5; DB 2; Length 149;
 Best Local Similarity 44.4%; Pred. No. 13;
 Matches 12; Conservative 7; Mismatches 5; Indels 3; Gaps 1;
 QY 3 ASVKYVILDMY---RALTLMTSTAT 26
 DB 104 APLAYVINDKFGIVGALITVHTTAT 130
 RESULT 12
 Q6D999 PRELIMINARY; PRT; 637 AA.
 AC Q6D999;
 DT 01-OCT-2004 (TReMBLrel. 28, Created)
 DT 01-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DT 01-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE ATP-independent RNA helicase.
 GN Name=dead; Synonyms=msb; ORFNames=ECA0718;
 OS Erwinia carotovora subsp. atroseptica SCR11043.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 OC NCBI_TaxID=218491;
 RN [1]_TaxID=218491;
 RP SEQUENCE FROM N.A.
 RC STRAIN=SCR11043;
 RA Bell K.S., Sebainia M., Pritchard L., Holden M., Hyman L.J.,
 RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
 RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
 RA Frazer A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
 RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
 RA Salmond G.P.C., Birch P.R.J., Barrall B.G., Parkhill J., Toth I.K.,
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; EX950851; CAG73633.1; -;
 KW Helicase.
 SQ SEQUENCE 637 AA; 70916 MW; 7CAC606D8B216B07 CRC64;
 Query Match 37.9%; Score 47; DB 2; Length 637;
 Best Local Similarity 50.9%; Pred. No. 76;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 10 LDMYRALLTLM 20
 DB 400 LDMYRALLTLM 410
 RESULT 13
 Q7XPT9 PRELIMINARY; PRT; 680 AA.
 AC Q7XPT9;
 DT 01-OCT-2003 (TReMBLrel. 25, Created)
 DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE OSJNBa008H09.17 protein.
 GN Names=OSJNBa008H09.17;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
 RA Liu Y.L., Ma J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L., Li T.,
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
 RA Zhang Y., Hu H., Jia P.X., Qian Y.W., Ying K., Zhou B., Chen Z.H.,
 RA Hae P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL606551; CAF03459.1; -;
 DR Gramene; Q7XPT9; -;
 DR GO; GO:0000145; C:exocyst; IEA.
 DR GO; GO:0006887; P:exocytosis; IEA.
 DR InterPro; IPR001917; Aminotrans_I.
 DR InterPro; IPR004140; Exo70.
 DR Pfam; PF03081; Exo70; 1.
 DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN 1.
 SQ SEQUENCE 680 AA; 77228 MW; 86A5D7F09F7464B9 CRC64;
 Query Match 37.9%; Score 47; DB 2; Length 680;
 Best Local Similarity 53.3%; Pred. No. 81;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 7 YVILDMYRALLTLMN 21
 DB 339 YILDMYKATLELQS 353
 RESULT 14
 Q86KT7 PRELIMINARY; PRT; 262 AA.
 AC Q86KT7;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
 OC NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA MEDLINE=22092622; PubMed=12097910;
 RA Gloeckner G., Eichinger L., Szafarski K., Pachebat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
 RL Nature 418:79-85 (2002).
 RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; AC115612; AAC0908.1; -;
KW Hypothetical protein.
SQ SEQUENCE 262 AA; 28290 MW; 5C5CECSBSA6FC987 CRC64;

Query Match 37.1%; Score 46; DB 2; Length 262;
Best Local Similarity 52.2%; Pred. No. 43;
Matches 12; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 QASVKYVILDMYRALTLMTST 24
| | | | | : | | | | |
Db 23 QESVHTVNIDNFPPLTLFNGST 45

RESULT 15

Q8PST9 PRELIMINARY; PRT; 641 AA.

AC Q8PST9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein XM2987.
GN OrderedLocusNames=XM2987;
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gcel / Gci / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierze A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR ENBL; AE013553; AM332683.1; -;
DR InterPro; IPR002035; VWF_A.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWA; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 641 AA; 73008 MW; E0D01942F3AEFF53 CRC64;

Query Match 37.1%; Score 46; DB 2; Length 641;
Best Local Similarity 43.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 EQASVKYVILDMYRAL 16
| | | | | : | | | | |
Db 134 EDARIEYILMDIYRGV 149

Search completed: November 10, 2004, 14:50:20
Job time : 36.634 secs

```

RESULT 2
US-08-261-663A-6
; Sequence 6, Application US/08261663A
; Patent No. 5571706
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
; APPLICANT: Whitham, Steven A
; TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Nicotiana glutinosa
US-09-813-742A-5

Query Match          44.7%; Score 46; DB 4; Length 652;
Best Local Similarity 36.8%; Pred. No. 41;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY      1 EDLESVLRLINWAKGSP1 19
DB      369 ENFEKLSLEVNYAKGLPL 387

RESULT 5
PCT-US95-07754A-6
; Sequence 6, Application PC/TUS9507754A
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
; APPLICANT: Whitham, Steven A
; TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07754A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0094.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07754A-6

Query Match          44.7%; Score 46; DB 5; Length 652;
Best Local Similarity 36.8%; Pred. No. 41;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY      1 EDLESVLRLINWAKGSP1 19
DB      369 ENFEKLSLEVNYAKGLPL 387

RESULT 6
US-08-310-912A-108
; Sequence 108, Application US/08310912A
; Patent No. S981730
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.

```

APPLICANT: Dahlbeck, Douglas
APPLICANT: Katagiri, Fumiaki
APPLICANT: Kunkel, Barbara N.
APPLICANT: Mindrinos, Michael N.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2904
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/310,912A
FILING DATE: September 22, 1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/227,360
FILING DATE: April 13, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Lech, Karen F.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/254001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 100254
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1143 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-310-912A-108

Query Match 44.7%; Score 46; DB 2; Length 1143;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRILINWAKGSP1 19
DB 369 ENFEKLSLEVNYAKGLPL 387

RESULT 7
US-09-301-085-108
Sequence 108, Application US/09301085
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick M.
APPLICANT: Staskawicz, Brian J.
APPLICANT: Brent, Andrew F.
APPLICANT: Dahlbeck, Douglas
APPLICANT: Katagiri, Fumiaki
APPLICANT: Kunkel, Barbara N.
APPLICANT: Mindrinos, Michael N.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
FILE REFERENCE: 00786/254002
CURRENT APPLICATION NUMBER: US/09/301,085
CURRENT FILING DATE: 1999-04-28
EARLIER APPLICATION NUMBER: 08/310,912
EARLIER FILING DATE: 1994-09-22
EARLIER APPLICATION NUMBER: 08/227,360

EARLIER FILING DATE: 1994-04-13
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 108
LENGTH: 1143
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-301-085-108

Query Match 44.7%; Score 46; DB 3; Length 1143;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRILINWAKGSP1 19
DB 369 ENFEKLSLEVNYAKGLPL 387

RESULT 8
PCT-US95-04589-108
Sequence 108, Application PC/TUS9504589
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick M.
APPLICANT: Staskawicz, Brian J.
APPLICANT: Brent, Andrew F.
APPLICANT: Dahlbeck, Douglas
APPLICANT: Katagiri, Fumiaki
APPLICANT: Kunkel, Barbara N.
APPLICANT: Mindrinos, Michael N.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: RPS2 GENE AND USES THEREOF
NUMBER OF SEQUENCES: 201
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2904
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04589
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,360
FILING DATE: 13-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/230001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 100254
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1143 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04589-108

Query Match 44.7%; Score 46; DB 5; Length 1143;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRILINWAKGSP1 19
DB 369 ENFEKLSLEVNYAKGLPL 387

Db 369 ENFEKLSLEVNYAKGLPL 387

RESULT 9

US-08-261-663A-2

Sequence 2, Application US/08261663A

Patent No. 5571706

GENERAL INFORMATION:

APPLICANT: Baker, Barbara J

APPLICANT: Whitham, Steven A

TITLE OF INVENTION: Plant Virus Resistance Gene and Methods

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Margaret A. Connor, USDA-ARS

STREET: 800 Buchanan Street

CITY: Albany

STATE: CA

COUNTRY: USA

ZIP: 94710

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/261,663A

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Connor, Margaret A

REGISTRATION NUMBER: 30043

REFERENCE/DOCKET NUMBER: 0094.94

TELEPHONE: (510) 559-5777

TELEFAX: (510) 559-5777

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1144 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-261-663A-2

Query Match 44.7%; Score 46; DB 1; Length 1144;

Best Local Similarity 36.8%; Pred. No. 77;

Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRLINWKGSP1 19

Db 369 ENFEKLSLEVNYAKGLPL 387

RESULT 10

US-08-261-663A-4

Sequence 4, Application US/08261663A

Patent No. 5571706

GENERAL INFORMATION:

APPLICANT: Baker, Barbara J

APPLICANT: Whitham, Steven A

TITLE OF INVENTION: Plant Virus Resistance Gene and Methods

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Margaret A. Connor, USDA-ARS

STREET: 800 Buchanan Street

CITY: Albany

STATE: CA

COUNTRY: USA

ZIP: 94710

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/261,663A

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/05272

FILING DATE: 15-APR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IL 113,373

FILING DATE: 13-APR-1995

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1144 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-261-663A-4

Query Match 44.7%; Score 46; DB 3; Length 1144;

Best Local Similarity 36.8%; Pred. No. 77;

Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRLINWKGSP1 19

Db 369 ENFEKLSLEVNYAKGLPL 387

RESULT 11

US-08-930-996A-9

Sequence 9, Application US/08930996A

Patent No. 6100449

GENERAL INFORMATION:

APPLICANT: FLUHR, Robert

APPLICANT: ESHED, Yuval

APPLICANT: ORI, Naomi

APPLICANT: PARAN, Ilan

APPLICANT: ZAMIR, Daniel

TITLE OF INVENTION: A GENE FAMILY FROM THE I2 FUSARIUM RESISTANCE

TITLE OF INVENTION: LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND

TITLE OF INVENTION: SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/930,996A

FILING DATE: 09-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/05272

FILING DATE: 15-APR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IL 113,373

FILING DATE: 13-APR-1995

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1144 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-930-996A-9

Query Match 44.7%; Score 46; DB 3; Length 1144;

Best Local Similarity 36.8%; Pred. No. 77;

Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRLINWKGSP1 19

Db 369 ENFEKLSLEVNYAKGLPL 387

RESULT 11

US-08-930-996A-9

Sequence 9, Application US/08930996A

Patent No. 6100449

GENERAL INFORMATION:

APPLICANT: FLUHR, Robert

APPLICANT: ESHED, Yuval

APPLICANT: ORI, Naomi

APPLICANT: PARAN, Ilan

APPLICANT: ZAMIR, Daniel

TITLE OF INVENTION: A GENE FAMILY FROM THE I2 FUSARIUM RESISTANCE

TITLE OF INVENTION: LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND

TITLE OF INVENTION: SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/930,996A

FILING DATE: 09-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/05272

FILING DATE: 15-APR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IL 113,373

FILING DATE: 13-APR-1995

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1144 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-930-996A-9

Query Match 44.7%; Score 46; DB 3; Length 1144;

Best Local Similarity 36.8%; Pred. No. 77;

Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRLINWKGSP1 19

Db 369 ENFEKLSLEVNYAKGLPL 387

```
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSP1 19
Db 369 ENFEKLSLEVVNYAKGLPL 387

RESULT 12
US-09-357-206A-3
; Sequence 3, Application US/09357206A
; Patent No. 6372962
; GENERAL INFORMATION:
; APPLICANT: Dinesh-Kumar, S.
; APPLICANT: Baker, Barbara
; TITLE OF INVENTION: Pathogen Resistance in Plants using cDNA-N/Intron Constructs
; FILE REFERENCE: 042250/191805 (5830-5)
; CURRENT APPLICATION NUMBER: US/09/357,206A
; CURRENT FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: US 60/093,494
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1144
; TYPE: PRT
; ORGANISM: Nicotiana glutinosa
US-09-357-206A-3

Query Match 44.7%; Score 46; DB 3; Length 1144;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSP1 19
Db 369 ENFEKLSLEVVNYAKGLPL 387

RESULT 13
US-09-813-742A-3
; Sequence 3, Application US/09813742A
; Patent No. 6630618
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
; APPLICANT: Dinesh-Kumar, S.P.
; TITLE OF INVENTION: NON-PATHOGEN INDUCED SYSTEMIC ACQUIRED RESISTANCE (SAR) IN PLANTS
; FILE REFERENCE: 42250/209601 (5830-12)
; CURRENT APPLICATION NUMBER: US/09/813,742A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,027
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1144
; TYPE: PRT
; ORGANISM: Nicotiana glutinosa
US-09-813-742A-3

Query Match 44.7%; Score 46; DB 4; Length 1144;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSP1 19
Db 369 ENFEKLSLEVVNYAKGLPL 387

RESULT 14
PCT-US95-07754A-2
; Sequence 2, Application PC/TUS9507754A
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
```

```
; APPLICANT: Whitham, Steven A
; TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07754A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0094.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07754A-2

Query Match 44.7%; Score 46; DB 5; Length 1144;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDLESVLRLINWAKGSP1 19
Db 369 ENFEKLSLEVVNYAKGLPL 387

RESULT 15
PCT-US95-07754A-4
; Sequence 4, Application PC/TUS9507754A
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara J
; APPLICANT: Whitham, Steven A
; TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07754A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0094.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067
```

```

; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07754A-4

Query Match 44.7%; Score 46; DS 5; Length 1144;
Best Local Similarity 36.8%; Pred. No. 77;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EDLESVLRLINWKGSP1 19
| : : : : : : : : : : : : : : : :
DB 369 ENFEKLSLEVNYAKGLPL 387

```

Search completed: November 10, 2004, 14:55:40
Job time : 7.37075 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 23.1321 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-56

Perfect score: 103

Sequence: 1 EDLESVLRLINWAKGSPIP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	100.0	20	14	US-10-092-750-56
2	50	48.5	226	16	US-10-767-701-44097
3	49	47.6	367	16	US-10-437-963-175166
4	49	47.6	393	16	US-10-437-963-175162
5	49	47.6	1288	14	US-10-114-824A-10
6	48	46.6	778	9	US-09-789-919-50
7	47	45.6	103	15	US-10-424-599-211799
8	47	45.6	108	15	US-10-424-599-265311
9	46	44.7	217	15	US-10-424-599-163276
10	46	44.7	423	14	US-10-369-493-8926
11	46	44.7	445	17	US-10-425-115-194216
12	46	44.7	652	9	US-09-813-742-5
13	46	44.7	1143	9	US-09-867-852-108

```

14 46 44.7 1143 15 US-10-613-472-108
15 46 44.7 1143 16 US-10-613-765-108
16 46 44.7 1144 9 US-09-813-742-3
17 45 43.7 72 15 US-10-424-599-212155
18 45 43.7 798 14 US-10-161-572-51
19 45 43.7 1163 14 US-10-295-027-1256
20 44.5 43.2 49 16 US-10-437-963-153840
21 44.5 43.2 425 15 US-10-425-114-68106
22 44.5 43.2 439 15 US-10-424-599-281104
23 44.5 43.2 1503 16 US-10-437-963-164301
24 44.5 43.2 1524 16 US-10-437-963-164299
25 44 42.7 84 16 US-10-437-963-138100
26 44 42.7 119 14 US-10-144-644-19
27 44 42.7 1125 16 US-10-700-740-19
28 44 42.7 125 17 US-10-425-115-285321
29 44 42.7 254 15 US-10-424-599-215229
30 44 42.7 311 15 US-10-425-114-59795
31 44 42.7 386 15 US-10-425-115-337442
32 44 42.7 401 17 US-10-425-115-337442
33 44 42.7 428 17 US-10-425-115-337443
34 44 42.7 430 15 US-10-389-566-2361
35 44 42.7 430 15 US-10-282-122A-47207
36 44 42.7 883 15 US-10-282-122A-70149
37 44 42.7 1551 14 US-10-318-906A-46
38 44 42.7 1551 14 US-10-319-236A-46
39 44 42.7 1551 14 US-10-418-036-12
40 44 42.7 1551 15 US-10-618-839-12
41 44 42.7 2780 14 US-10-220-587-2
42 44 42.7 2780 14 US-10-423-483-2
43 43.5 42.2 150 15 US-10-424-599-218873
44 43 41.7 49 14 US-10-029-386-31646
45 43 41.7 52 9 US-09-832-197-7

```

ALIGNMENTS

```

RESULT 1
US-10-092-750-56
; Sequence 56, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-56

```

Query Match 100.0%; Score 103; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 EDLESVLRLINWAKGSPIP 20
    |||||
Db 1 EDLESVLRLINWAKGSPIP 20
    |||||

```

```

RESULT 2
US-10-767-701-44097
; Sequence 44097, Application US/10767701
; Publication No. US2004012684A1
; GENERAL INFORMATION:

```

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 44097
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(226)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C20537_1.pep
US-10-767-701-44097

Query Match 48.5%; Score 50; DB 16; Length 226;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 LESVLRLINWAKGSPIP 20
Db 141 VEEVLKDDIDWLPGSPPEP 158

RESULT 3
US-10-437-963-175166
; Sequence 175166, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 175166
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRP4530_73038C.1.pep
US-10-437-963-175166

Query Match 47.6%; Score 49; DB 16; Length 367;
Best Local Similarity 46.7%; Pred. No. 28;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 VLIRLINWAKGSPIP 20
Db 279 IVCRLISWCKTAFVP 293

RESULT 4
US-10-437-963-175162
; Sequence 175162, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 175162
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(393)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRP4530_73034C.1.pep
US-10-437-963-175162

Query Match 47.6%; Score 49; DB 16; Length 393;
Best Local Similarity 46.7%; Pred. No. 31;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 VLIRLINWAKGSPIP 20
Db 305 IVCRLISWCKTAFVP 319

RESULT 5
US-10-114-824A-10
; Sequence 10, Application US/10114824A
; Publication No. US20030196215A1
; GENERAL INFORMATION:
; APPLICANT: JOSELYNE OLIVIER
; TITLE OF INVENTION: Pathogenic Agents
; FILE REFERENCE: CHEP:006US
; CURRENT APPLICATION NUMBER: US/10/114,824A
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1288
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-114-824A-10

Query Match 47.6%; Score 49; DB 14; Length 1288;
Best Local Similarity 36.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDLESVLRLINWAKGSP 19
Db 324 QNLQELSVRVINYANGNPL 342

RESULT 6
US-09-789-919-50
; Sequence 50, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; APPLICANT: Moore, Kateri
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; TITLE OF INVENTION: CELLS AND USES THEREOF
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21

```

; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: All Xaa's in this sequence represent unreadable
; OTHER INFORMATION: amino acid.
US-09-789-919-50

Query Match          46.6%; Score 48; DB 9; Length 778;
Best Local Similarity 47.1%; Pred. No. 94;
Matches             8; Conservative 5; Mismatches 0; Gaps 0;

QY      2 DLESLVRLINWKGSP 18
      ||::||:::||
Db      472 DLQFLINIVFCKNSP 488

RESULT 7
US-10-424-599-211799
; Sequence 211799, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211799
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33281C.1.pep
US-10-424-599-211799

Query Match          45.6%; Score 47; DB 15; Length 103;
Best Local Similarity 70.0%; Pred. No. 15;
Matches             7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      9 RLINWAKGSP 18
      ||::||:::||
Db      84 RVLNWKGP 93

RESULT 8
US-10-424-599-265311
; Sequence 265311, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265311
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT3847 81596C.1.pep

```

```

US-10-424-599-265311

Query Match      45.6%; Score 47; DB 15; Length 108;
Best Local Similarity 57.1%; Pred. No. 15;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;

QY      6 VLIRLINWAKGSP I 19
      ||:|||||:|:|:|
DB      86 VLRLRNWTKDPL 99

RESULT 9
US-10-424-599-163276
; Sequence 163276, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 163276
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118457C.1.pap
US-10-424-599-163276

Query Match      44.7%; Score 46; DB 15; Length 217;
Best Local Similarity 44.4%; Pred. No. 48;
Matches      8; Conservative      5; Mismatches      5; Indels      0; Gaps      0;

QY      3 LBSVLIRLINWAKGSP I 20
      ||:||:||||:|:|
DB      119 MEKVOKLLNWAKEAGLP 136

RESULT 10
US-10-369-493-8926
; Sequence 8926, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8926
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Chloroflexus aurantiacus
US-10-369-493-8926

Query Match      44.7%; Score 46; DB 14; Length 423;
Best Local Similarity 56.2%; Pred. No. 1e+02;
Matches      9; Conservative      4; Mismatches      3; Indels      0; Gaps      0;

QY      1 EDLESVLIRLINWAKG 16
      ||:||:|||||

```

```

; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: All Xaa's in this sequence represent unreadable
; OTHER INFORMATION: amino acid.
US-09-789-919-50

Query Match          46.6%; Score 48; DB 9; Length 778;
Best Local Similarity 47.1%; Pred. No. 94;
Matches             8; Conservative 5; Mismatches 0; Gaps 0;

QY      2 DLESLVRLINWKGSP 18
      ||::||:::||
Db      472 DLQFLINIVFCKNSP 488

RESULT 7
US-10-424-599-211799
; Sequence 211799, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211799
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33281C.1.pep
US-10-424-599-211799

Query Match          45.6%; Score 47; DB 15; Length 103;
Best Local Similarity 70.0%; Pred. No. 15;
Matches             7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      9 RLINWAKGSP 18
      ||::||:::||
Db      84 RVLNWKGP 93

RESULT 8
US-10-424-599-265311
; Sequence 265311, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265311
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT3847 81596C.1.pep

```

```
Db      162 EDVENILLRLQAADG 177

RESULT 11
US-10-425-115-194216
; Sequence 194216, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 194216
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Zea mays.
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(445)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_108705C.1.pap
US-10-425-115-194216
Query Match      44.7%; Score 46; DB 17; Length 445;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy      3 LSVLRLINWAKGSP 20
      : : : : :
Db      142 VEEFVKIDWLPKSGPE 159

RESULT 12
US-09-813-742-5
; Sequence 5, Application US/09813742
; Patent No. US20020004944A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Barbara
; APPLICANT: Dinesh-Kumar, S.P.
; TITLE OF INVENTION: NON-PATHOGEN INDUCED SYSTEMIC ACQUIRED RESISTANCE (SAR) IN PLANTS
; FILE REFERENCE: 042250/209601 (S830-12)
; CURRENT APPLICATION NUMBER: US/09/813,742
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,027
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-09-813-742-5
Query Match      44.7%; Score 46; DB 9; Length 652;
Best Local Similarity 36.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy      1 EDLSVLRLINWAKGSP 19
      : : : : :
Db      369 ENFEKLSLEVNYAKGLPL 387

RESULT 13
US-09-867-852-108
; Sequence 108, Application US/09867852
; Patent No. US20020147324A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
; FILE REFERENCE: 00786/254002
; CURRENT APPLICATION NUMBER: US/09/867,852
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/301,085
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/310,912
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-09-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/227,360
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-867-852-108
Query Match      44.7%; Score 46; DB 9; Length 1143;
Best Local Similarity 36.8%; Pred. No. 3e+02;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy      1 EDLSVLRLINWAKGSP 19
      : : : : :
Db      369 ENFEKLSLEVNYAKGLPL 387

RESULT 14
US-10-613-472-108
; Sequence 108, Application US/10613472
; Publication No. US20040088756A1
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
; FILE REFERENCE: 00786/254004
; CURRENT APPLICATION NUMBER: US/10/613,472
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 09/867,852
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 09/301,085
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: US 08/310,912
; PRIOR FILING DATE: 1994-09-22
; PRIOR APPLICATION NUMBER: US 08/227,360
; PRIOR FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-613-472-108
Query Match      44.7%; Score 46; DB 15; Length 1143;
Best Local Similarity 36.8%; Pred. No. 3e+02;
```

Matches	7;	Conservative	7;	Mismatches	5;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

```
QY      1 EDLESVLIRLINWAKGSP I 19
         || : | : : ||| :
Db     369 ENFEKLSLEVWVNYAKGLPL 387
```

RESULT 15

```

US-10-613-765-108
; Sequence 108, Application US/10613765
; Publication No. US20040172673A1
; GENERAL INFORMATION:
; APPLICANT: AUSUBEL, FREDERICK M.
; APPLICANT: STASKAWICZ, BRIAN J.
; APPLICANT: KATAGIRI, FUMIAKI
; APPLICANT: KUNKEL, BARBARA N.
; APPLICANT: MINDRINOS, MICHAEL N.
; APPLICANT: YU, GUO-LIANG
; APPLICANT: BAKER, BARBARA
; APPLICANT: ELLIS, JEFFREY
; APPLICANT: SALMERON, JOHN
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, AND
; TITLE OF INVENTION: DETECTION METHODS
; FILE REFERENCE: 00786/254005
; CURRENT APPLICATION NUMBER: US/10/613.765
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 09/867,852
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 09/301,085
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: US 08/310,912
; PRIOR FILING DATE: 1994-09-22
; PRIOR APPLICATION NUMBER: US 08/227,360
; PRIOR FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-613-765-108

```

```
QY      1 EDLESVLIRLINWAXGSP I 19
        | : | : | : | : | : | :
Dp     369 ENFEKLSLEVNYAKGLPL 387
```

Search completed: November 11, 2004, 02:43:12
Job time : 24.1821 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 4.90566 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-56

Perfect score: 103

Sequence: 1 EDLSVLRLINWAKGSPIP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	48.5	392	B32370	cyclin B2 - Africa
2	49	47.6	935	E96806	hypothetical prote
3	48	46.6	20	S59382	hypothetical prote
4	48	46.6	1647	T41267	hypothetical prote
5	47	45.6	848	JG0194	androgen receptor
6	46	44.7	756	T48139	copper amine oxida
7	46	44.7	1144	A54810	TMV resistance pro
8	46	44.7	1186	H88869	protein unc-31 (im
9	45	43.7	250	E72605	probable high-affi
10	45	43.7	297	F70201	conserved hypotet
11	45	43.7	381	D89922	conserved hypotet
12	45	43.7	397	A32370	cyclin B1 - Africa
13	45	43.7	431	T29267	hypothetical prote
14	45	43.7	611	T41563	hypothetical prote
15	45	43.7	1522	JC1101	sodium channel pro
16	44.5	43.2	170	T08821	disease resistance
17	44.5	43.2	326	AF2024	hypothetical prote
18	44.5	43.2	414	A72765	hypothetical prote
19	44.5	43.2	583	S19476	hypothetical prote
20	44	42.7	244	1 LNRTMC	mannose-binding le
21	44	42.7	290	A23519	isochorismatase (E
22	44	42.7	394	T42512	1D-myo-inositol-tr
23	44	42.7	429	AD3642	(S)-2-hydroxy-acid
24	44	42.7	430	C70176	probable ATP-depen
25	44	42.7	461	T42513	1D-myo-inositol-tr
26	44	42.7	486	T25639	hypothetical prote
27	44	42.7	494	T42444	1D-myo-inositol-tr
28	44	42.7	648	1 DJSPS2	DNA-directed DNA p
29	44	42.7	725	T35114	probable kinase/ph

MRP-like ABC trans
conserved hypotet
hypothetical prote
hypothetical prote
cyclin B - rat
cyclin B - long-ta
cyclin B1 - mouse
cyclin B1 - mouse
SMV2 protein - Yea
cobalamin biosynth
glutamate-ammonia-
hypothetical prote
mannose-binding le
lactose transport
hypothetical prote

ALIGNMENTS

RESULT 1

B32370

cyclin B2 - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 02-Nov-1989 #sequence_revision 02-Nov-1989 #text_change 12-Jul-2004

C;Accession: B32370

R;Minshull, J.; Blow, J.J.; Hunt, T.

Cell 56, 947-956, 1989

A;Title: Translation of cyclin mRNA is necessary for extracts of activated Xenopus eggs

A;Reference number: A32370; MUID:89168446; PMID:2564315

A;Accession: B32370

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-392 <MIN>

A;Cross-references: UNIPROT:P13351; GB:J03167; NID:G214094; PIDN:AAA49697.1; PID:G21409

C;Superfamily: cyclin, A/B/D/E type

C;Keywords: cell cycle control

Query Match 48.5%; Score 50; DB 2; Length 392;

Best Local Similarity 36.8%; Pred. No. 4.1;

Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLBSVLRLINWAKGSPIP 20

DB 245 EMEMILRLNFDLGRPLP 263

RESULT 2

E96806

hypothetical protein T32E8.1 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: E96806

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: E96806

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-935 <SFC>

A;Cross-references: UNIPROT:Q9CA26; GB:AB005173; NID:96437530; PIDN:AAF08562.1; GSPDB:G

C;Genetics:

A;Gene: T32E8.1

A;Map position: 1

Db | : | : | : | : ||| : |
369 ENFEKLSLEVYVYAKGLPL 387

RESULT 8
H88569
protein unc-31 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: H88869
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:U9069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: H88869
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1186 <STO>
A;Cross-references: UNIPROT:Q23658; GB:chr_IV; PIDN:CAA93520.1; PID:g3881885; GSPDB:GN00
C;Genetics:
A;Gene: unc-31
A;Map position: 4
C;Superfamily: calcium-dependent actin-binding protein; pleckstrin repeat homology

Query Match 44.7%; Score 46; DB 2; Length 1186;
Best Local Similarity 58.8%; Pred. No. 56;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LESVLRLINWAGSPI 19
||||| : |||
Db 1154 LESVLSRLARYDGNPI 1170

RESULT 9
E72605
probable high-affinity branched-chain amino acid transport ATP-binding protein APEI308 -
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 16-Aug-2004
C;Accession: E72605
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; N
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: E72605
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-250 <KAW>
A;Cross-references: UNIPROT:Q9YCE9; DDBJ:AP000061; NID:g5104821; PIDN:BAA80299.1; PID:d
A;Experimental source: strain K1
C;Genetics:
A;Gene: APEI308
C;Superfamily: ATP-binding cassette homology

Query Match 43.7%; Score 45; DB 2; Length 250;
Best Local Similarity 46.7%; Pred. No. 16;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LESVLRLINWAGKS 17
||| : |||
Db 94 LENIMVRLHPWTGRS 108

RESULT 10
F70201
conserved hypothetical protein BB0815 - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C;Accession: F70201
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavang, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt
son, D.; Peterson, J.; Kerlavang, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 26.6415 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-56

Perfect score: 103
Sequence: 1 EDLESVLIRLINWAKGSPIP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	49.5	471	2 Q88GL4	P88914 pseudomonas
2	50	48.5	392	1 Q88GL4	P13351 xenopus lae
3	50	48.5	761	1 AD24_MOUSE	Q9r160 mus musculu
4	50	48.5	761	2 Q8CDV3	Q8cdv3 mus musculu
5	49	47.6	133	2 Q7ND57	Q7nd57 Gloeobacter
6	49	47.6	935	2 Q9CA26	Q9ca26 arabidopsis
7	49	47.6	1187	2 Q8GZ83	Q8gz83 arabidopsis
8	49	47.6	1187	2 Q9FH84	Q9fh84 arabidopsis
9	48.5	47.1	150	2 Q84Y05	Q84y05 glycine max
10	48	46.6	594	2 Q745W1	Q745w1 thermus the
11	48	46.6	594	2 AA882524	AA882524 thermus t
12	48	46.6	760	2 Q8X1N6	Q8x1n6 cephalospor
13	48	46.6	1104	2 Q922R7	Q922r7 rattus norv
14	48	46.6	1647	2 P78847	P78847 schizosacch
15	47.5	46.1	554	2 Q9SF12	Q9sf12 arabidopsis
16	47	45.6	137	2 Q81H29	Q81h29 bacillus ce
17	47	45.6	147	2 Q8UWF7	Q8uwf7 salmo salar
18	47	45.6	166	2 Q9DDA0	Q9dda0 xenopus tro
19	47	45.6	392	2 Q6FA39	Q6fa39 xenopus lae
20	47	45.6	392	2 AAH60466	AAH60466 xenopus l
21	47	45.6	743	2 Q97120	Q97120 schistosoma
22	47	45.6	771	2 Q800S7	Q800s7 acanthopagr
23	47	45.6	848	2 Q9YGV9	Q9ygv9 anguilla ja
24	47	45.6	853	2 Q93245	Q93245 oncorhynch
25	46	44.7	128	2 Q9PUL1	Q9pull cryzias lat
26	46	44.7	130	2 Q9OWA3	Q9owa3 spnoeroides
27	46	44.7	200	2 Q918F5	Q918f5 pimephales
28	46	44.7	310	2 Q73109	Q73109 wlbachia p
29	46	44.7	310	2 AA514103	AA514103 wlbachia
30	46	44.7	362	2 Q9T9H3	Q9t9h3 halocynthia
31	46	44.7	377	2 Q6EAL5	Q6eal5 brachydanio

ALIGNMENTS

RESULT 1

Q88GL4 ID Q88GL4 PRELIMINARY; PRT; 471 AA.
AC Q88GL4; (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=PP3707;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzaz A., Uytterback T.R., Rizzo W., Lee K., Kosack D., Woelfel D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnsels J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duisterhoef A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AF016788; AA69304.1; -.
DR TIGR; PP3707; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 471 AA; 52925 MW; 20A9B3CE128BDEF5 CRC64;

Query Match 49.5%; Score 51; DB 2; Length 471;
Best Local Similarity 42.1%; Pred. No. 16;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EDLESVLIRLINWAKGSPIP 19

Db 449 EAMEKLIKQDMLGSPV 467

RESULT 2

Q88GL4 ID Q88GL4 PRELIMINARY; PRT; 471 AA.
AC Q88GL4; (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=PP3707;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzaz A., Uytterback T.R., Rizzo W., Lee K., Kosack D., Woelfel D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnsels J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duisterhoef A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AF016788; AA69304.1; -.
DR TIGR; PP3707; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 471 AA; 52925 MW; 20A9B3CE128BDEF5 CRC64;

Query Match 49.5%; Score 51; DB 2; Length 471;
Best Local Similarity 42.1%; Pred. No. 16;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EDLESVLIRLINWAKGSPIP 19

Db 449 EAMEKLIKQDMLGSPV 467

RX NCBI_TaxID=8355;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=89168446; PubMed=2564315;
 RX Minshull J., Blow J.J., Hunt T.;
 RA "Translation of cyclin mRNA is necessary for extracts of activated
 RT xenopus eggs to enter mitosis.";
 RL Cell 56:947-956(1989).
 CC -!- FUNCTION: Essential for the control of the cell cycle at the G2/M
 CC (mitosis) transition.
 CC -!- SUBUNIT: Interacts with the CDC2 protein kinase to form a
 CC serine/threonine kinase holoenzyme complex also known as
 CC maturation promoting factor (MPF). The cyclin subunit imparts
 CC substrate specificity to the complex.
 CC -!- DEVELOPMENTAL STAGE: Accumulates steadily during G2 and is
 CC abruptly destroyed at mitosis.
 CC -!- SIMILARITY: Belongs to the cyclin family. Cyclin AB subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; J03167; AAA49697.1; -.
 DR PIR; B32370; B32370.
 DR HSSP; P20248; 1H1R.
 DR InterPro; IPR006670; Cyclin.
 DR InterPro; IPR004367; Cyclin_Cterm.
 DR InterPro; IPR011028; Cyclin_Like.
 DR InterPro; IPR006671; Cyclin_N.
 DR Pfam; PF02984; Cyclin_C; 1.
 DR Pfam; PF00134; Cyclin_N; 1.
 DR SMART; SM00385; CYCLIN; 2.
 DR PROSITE; PS00292; CYCLIN; 1.
 KW Cell cycle; Cell division; Cyclin; Mitosis.
 SQ SEQUENCE 392 AA; 43624 MW; 87955E04C106218E CRC64;

 Query Match 48.5%; Score 50; DB 1; Length 392;
 Best Local Similarity 36.8%; Pred. No. 19;
 Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

 QY 2 DLESLVRLINNAKGSPIP 20
 Db 245 EMEMILRLNFDLGRPLP 263

 RESULT 3
 AD24_MOUSE STANDARD; PRT; 761 AA.
 ID AD24_MOUSE
 AC Q9R160;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE ADAM 24 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE domain 24) (Testase 1).
 GN Name-Adam24;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Testis;
 RC MEDLINE=99326496; PubMed=10395895;
 RX Zhu G.-Z., Lin Y., Myles D.G., Primakoff P.;
 RA "Identification of four novel ADAMS with potential roles in
 RT spermatogenesis and fertilization.";
 RL Gene 234:227-237(1999).
 RN [2]
 RP CHARACTERIZATION.

RX MEDLINE=21206048; PubMed=11309208;
 RA Zhu G.-Z., Myles D.G., Primakoff P.;
 RT "Testase 1 (ADAM 24) a plasma membrane-anchored sperm protease
 RT implicated in sperm function during epididymal maturation or
 RL J. Cell Sci. 114:1787-1794(2001).
 CC -!- FUNCTION: Plasma membrane protease present on mature sperm that
 CC may be involved in sperm function during epididymal maturation
 CC and/or fertilization.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed exclusively in testis and more
 CC specifically on the surface of mature sperm. Localized to the
 CC equatorial region of the plasma membrane of cauda epididymal
 CC sperm.
 CC -!- DEVELOPMENTAL STAGE: Adult levels are reached by day 20 after
 CC birth.
 CC -!- PFM: The prodomain is removed during sperm passage through the
 CC caput epididymis after the protein has reached the cell surface.
 CC Not processed in the secretory pathway.
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF167402; AAD48841.1; -.
 DR HSSP; P18619; IFVL.
 DR MEROPS; M12.227; -.
 DR MGD; MGI:105984; Adam24.
 DR InterPro; IPR006586; ADAM cysteine.
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR002870; Peptidase_M12B_N.
 DR InterPro; IPR001818; Pept_M12B_M12B_N.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR Pfam; PF00200; Disintegrin; 1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR ProDom; PD000664; Disintegrin; 1.
 DR SMART; SM00608; ACR; 1.
 DR SMART; SM00050; DISIN; 1.
 DR PROSITE; PS50215; ADAM_MBP; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS50214; DISINTEGRIN_2; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Metalloprotease; Signal;
 KW Spermatogenesis; Transmembrane; Zinc; Zymogen.
 KW SIGNAL 1 34 Potential.
 FT PROPEP 35 200 By similarity.
 FT CHAIN 201 761 ADAM 24.
 FT DOMAIN 35 697 Extracellular (Potential).
 FT TRANSMEM 698 718 Potential.
 FT DOMAIN 719 761 Cytoplasmic (Potential).
 FT DOMAIN 201 405 Metalloprotease.
 FT DOMAIN 406 493 Disintegrin-like.
 FT DOMAIN 494 630 Cys-rich.
 FT DOMAIN 631 664 EGF-like.
 FT SITE 174 174 Cysteine switch (Potential).
 FT METAL 342 342 Zinc (catalytic) (By similarity).

Page 4

```

DR InterPro; IPRO00157; TTR.
DR Pfam; PF00560; LRR; 5.
DR Pfam; PF00931; NB-ARC; 2.
DR PRINTS; PRC0364; DISEASERIST.
SQ SEQUENCE 1187 AA; 134275 MW; 7EF6A2240C692B11 CRC64;

Query Match 47.6%; Score 49; DB 2; Length 1187;
Best Local Similarity 36.8%; Pred.No.99;
Matches 7; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDLESVLIRLINWAKGSP1 19
   :|: : |||||:|:
DB 324 QNLQSLSVRVINYANGNPL 342

RESULT 9
Q84Y05 PRELIMINARY; PRT; 150 AA.
AC Q84Y05;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Disease resistance protein-like protein (Fragment).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu C., Santos F.A., Nimmakayala P., Springman R., Meksem K.,
RL Lightfoot D.A., Zhang H.-B.;
RA Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY170485; AAC18343.1; -;
DR GO; GO:0005224; P:ATP binding; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
FT NON TER 1
SQ SEQUENCE 150 AA; 17391 MW; 653B831035765D57 CRC64;

Query Match 47.1%; Score 48.5; DB 2; Length 150;
Best Local Similarity 42.3%; Pred.No.11;
Matches 11; Conservative 3; Mismatches 5; Indels 7; Gaps 1;

QY 1 EDLESVL-----IRLINWAKGSP1 19
   |||||
DB 29 EDLSQAFSLKTAGEIKLTNWREGIPI 54
   |||||

RESULT 10
Q745W1 PRELIMINARY; PRT; 594 AA.
AC Q745W1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=TTPO194;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
CG Plasmid pT27.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;
RN [1]
RP SEQUENCE FROM N.A.
RA Henne A., Brueggemann H., Raasch C., Wierer A., Hartsch T.,
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA Jacobi C., Starkuviene V., Schlenczek S., Dencker S., Huber R.,
RA Klein H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT the genome sequence of the extreme thermophile Thermus
thermophilus.";
```

```

RA Abe I., Naito K., Takagi Y., Noguchi H.;
RT "Molecular cloning, expression, and site-directed mutations of
RT Oxidoreductase cyclase from Cephalosporium caerulens.";
RL Biochim. Biophys. Acta 1522:67-73(2001).
DR EMBL: AF327881; AAL56020.1; -.
DR GO: GO:001629; F:lyase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001330; P:prenyltrans.
DR InterPro: IPR002365; Terpen synth.
DR InterPro: IPR008930; Terp cyc toroid.
DR Pfam: PF00432; P:prenyltrans; 3_.
DR TIGRfam: TIGR01787; squalene_cyclase; 1.
SQ SEQUENCE 760 AA; 87082 MW; 49C01F4B50D15FE1 CRC64;

Query Match 46.6%; Score 48; DB 2; Length 760;
Best Local Similarity 42.1%; Pred. No. 87;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 EDLESVLIRLINWAKSP1 19
DB 579 EDIEKFIKRNWIRGSSL 597

RESULT 13
ID Q922R7 PRELIMINARY; PRT; 1104 AA.
AC Q922R7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Replication factor C (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1
RP SEQUENCE FROM N.A.
RA Blum A., Simon B.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF303050; AD01890.1; -.
DR GO: GO:0005663; C:DNA replication factor C complex; IEA.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006260; P:DNA replication; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003959; AAA_ATPase_cent.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR000862; RFC.
DR Pfam: PF00004; AAA; 1.
DR Pfam: PF00533; BRCT; 1.
DR SMART: SM00382; AAA; 1.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS00172; BRCT; 1.
KW ATP-binding.
FT NON TER 1104 1104
SQ SEQUENCE 1104 AA; 123246 MW; B8E1E4BCE650793B CRC64;

Query Match 46.6%; Score 48; DB 2; Length 1104;
Best Local Similarity 45.0%; Pred. No. 13e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 EDLESVLIRLINWAKSP1 20
DB 87 EKLPRVLNVLNLSGSSLP 106

RESULT 14
P78847
ID P78847 PRELIMINARY; PRT; 1647 AA.
AC P78847;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)

```

```

DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SPC290.03C protein.
GN Name=SPC290.03C;
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN 1
RP SEQUENCE FROM N.A.
STRAIN=972h-;
MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgours J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welter J., Vanstreels E., Rieger M., Schafer M., Muller-Auer S.,
RA Gabel C., Fuchs M., Dusterhoft A., Fritz C., Holzer E., Moestl D.,
RA Hilbert H., Borzym K., Langer I., Beck A., Leinrich H., Reinhardt R.,
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas R., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
DR EMBL: AL035260; CAA22873.1; -.
DR PIR: T41267;
DR PIR: T42996; T42996.
DR GeneDB Spombe; SPC290.03C; -.
DR GO: GO:0005635; C:nuclear membrane; IEA.
DR GO: GO:0000917; P:barrier septum formation; IMP.
SQ SEQUENCE 1647 AA; 186410 MW; 579B36CB585ED618 CRC64;

Query Match 46.6%; Score 48; DB 2; Length 1647;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 EDLESVLIRLINWAKSP1 20
DB 477 DDIESDMVGLTWSMGSLP 496

RESULT 15
Q9SF12
ID Q9SF12 PRELIMINARY; PRT; 554 AA.
AC Q9SF12;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F26K24.16.
GN Name=F26K24.16;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 1
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,

```



```

RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC016795; AAF23203.1; -
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR006567; PUG.
DR InterPro; IPR010513; Ribonuc_2-5A.
DR Pfam; PF00069; Kinase; 1.
DR Pfam; PF06479; Ribonuc_2-5A; 1.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SM00580; PUG; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein.
SQ SEQUENCE 554 AA; 63259 MW; 3543FFF1A6651683 CRC64;

Query Match 46.1%; Score 47.5; DB 2; Length 554;
Best Local Similarity 47.6%; Pred. NC. 73;
Matches 10; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 EDLESVLRLNW-AKGSPIP 20
Db 224 DDLEKVMKRIKFWKEGRPLP 244

Search completed: November 10, 2004, 14:50:22
Job time : 28.6415 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 9.15094 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-57

Perfect score: 137

Sequence: 1 RPVSFCGAVWTLNRAIGHFVRGSR 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	35.8	376	1	US-08-002-024B-8
2	48	35.0	919	4	US-08-583-681A-7187
3	47	34.3	995	4	US-08-252-991A-22297
4	46	33.6	464	4	US-08-252-991A-27800
5	45	32.8	252	4	US-08-252-991A-26127
6	45	32.8	452	4	US-08-489-039A-9355
7	45	32.8	607	3	US-08-969-683A-59
8	45	32.8	1693	3	US-08-478-507-7
9	45	32.8	1693	3	US-08-128-275A-7
10	44.5	32.8	1693	3	US-09-553-427-7
11	44.5	32.5	107	4	US-09-673-395A-392
12	44.5	32.5	175	4	US-08-270-767-41770
13	44	32.1	63	2	US-08-637-759B-442
14	44	32.1	63	3	US-08-871-355A-442
15	44	32.1	63	3	US-09-201-945-442
16	44	32.1	235	4	US-08-489-039A-7572
17	44	32.1	237	4	US-08-489-039A-9365
18	44	32.1	400	4	US-09-134-000C-4389
19	44	32.1	1708	4	US-09-462-606-2
20	43.5	31.8	378	2	US-08-846-762-87
21	43.5	31.8	405	4	US-09-489-039A-9838
22	43	31.4	371	3	US-09-133-069-2
23	43	31.4	455	4	US-08-248-796A-16498
24	43	31.4	528	4	US-10-140-372-7
25	43	31.4	730	2	US-08-696-944-2
26	42.5	31.0	206	4	US-09-270-767-57797
27	42.5	31.0	274	4	US-09-270-767-42500

28	42.5	31.0	470	4	US-09-248-796A-15129	Sequence 15129, A
29	42.5	31.0	799	4	US-09-489-039A-8219	Sequence 8219, Ap
30	42	30.7	59	4	US-09-513-999C-6346	Sequence 6346, Ap
31	42	30.7	60	4	US-09-540-236-3288	Sequence 3288, Ap
32	42	30.7	181	4	US-09-270-767-40764	Sequence 40764, A
33	42	30.7	161	4	US-09-270-767-55980	Sequence 55980, A
34	42	30.7	162	4	US-09-252-991A-26062	Sequence 26062, A
35	42	30.7	184	4	US-09-252-991A-28524	Sequence 28524, A
36	42	30.7	460	4	US-09-489-039A-12680	Sequence 12680, A
37	42	30.7	477	1	US-07-735-065-2	Sequence 2, Appli
38	42	30.7	477	1	US-08-469-202-12	Sequence 12, Appli
39	42	30.7	477	2	US-08-484-434C-12	Sequence 12, Appli
40	42	30.7	477	4	US-09-384-361-12	Sequence 12, Appli
41	42	30.7	580	3	US-08-818-112-75	Sequence 75, Appli
42	42	30.7	580	3	US-08-818-111-76	Sequence 76, Appli
43	42	30.7	580	3	US-09-056-556-75	Sequence 75, Appli
44	42	30.7	580	4	US-09-072-596-76	Sequence 76, Appli
45	42	30.7	580	4	US-09-072-967-75	Sequence 75, Appli

ALIGNMENTS

RESULT 1
US-08-002-024B-8
; Sequence 8, Application US/08002024B
; Patent No. 5798103
; GENERAL INFORMATION:
; APPLICANT: MOOI, Frederik R
; TITLE OF INVENTION: WHOOPING COUGH VACCINE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: VA
; COUNTRY: US
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/002,024B
; FILING DATE: 08-JAN-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92200038.5
; FILING DATE: 08-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Robert J
; REGISTRATION NUMBER: 17,355
; REFERENCE/DOCKET NUMBER: BO-37424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703/521-2297
; TELEFAX: 703/685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-002-024B-8

Query Match 35.8%; Score 49; DB 1; Length 376;
Best Local Similarity 34.2%; Pred. No. 9.1;
Matches 13; Conservative 2; Mismatches 5; Indels 18; Gaps 2;

QY 1 RPVSFCG---AVMT-----LNRAIGHF 20
DB 15 RVISFCGALAWAGLAVQPMANAVDPVDCGALGLHF 52

```
RESULT 2
US-09-543-681A-7187
; Sequence 7187, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7187
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7187

Query Match 35.0%; Score 48; DB 4; Length 919;
Best Local Similarity 41.7%; Pred. No. 37;
Matches 10; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 1 RPVSFCGAVWTLNRAIGHRVGRS 24
|||:||||: : |||:
Db 494 RPVAFCGG---TSVAGYIVRG 513

RESULT 3
US-09-252-991A-22297
; Sequence 22297, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22297
; LENGTH: 995
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22297

Query Match 34.3%; Score 47; DB 4; Length 995;
Best Local Similarity 52.6%; Pred. No. 58;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 4 SFCGAVWTLNRAIGHRVGR 22
|||:||||: : |||:
Db 160 SFGALWLFERAVRAGR 178

RESULT 4
US-09-252-991A-27800
; Sequence 27800, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27800
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27800

Query Match 33.6%; Score 46; DB 4; Length 464;
Best Local Similarity 52.6%; Pred. No. 35;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 PVSFCGAVWTLNRAIGHRF 20
|||:||||: : |||:
Db 115 PVLATGAATYLNALGTHF 133

RESULT 5
US-09-252-991A-26127
; Sequence 26127, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26127
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26127

Query Match 32.8%; Score 45; DB 4; Length 252;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 14 RAIGHRVGRS 25
|||:||||: : |||:
Db 80 RATGRHWARGTR 91

RESULT 6
US-09-489-039A-9355
; Sequence 9355, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9355
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9355

Query Match 32.8%; Score 45; DB 4; Length 452;
Best Local Similarity 56.2%; Pred. No. 49;
```

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSFCGAVWTLNRAICR 18
DB 145 VDFLDVWVWLNALDR 160

RESULT 7
US-08-969-683A-59
; Sequence 59, Application US/08969683A
; Patent No. 6136576
; GENERAL INFORMATION:
; APPLICANT: GENENCOR INTERNATIONAL, INC.
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; TITLE OF INVENTION: PRODUCTION OF 1,3 PROPANEDIOL
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 4 Cambridge Place
; STREET: 1870 South Wincon road
; CITY: Rochester
; STATE: NY
; COUNTRY: U.S.A
; ZIP: 14618
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,683A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/20873
; FILING DATE: 13-NOV-1997
; APPLICATION NUMBER: 60/030,601
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC 369-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-864-7620
; TELEFAX: 650-845-6504
; TELEX:
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 607 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6136576e
US-08-969-683A-59

Query Match 32.8%; Score 45; DB 3; Length 607;
Best Local Similarity 56.2%; Pred. No. 69;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSFCGAVWTLNRAICR 18
DB 141 VDFLDVWVWLNALDR 156

RESULT 8
US-08-478-507-7
; Sequence 7, Application US/08478507
; Patent No. 6120988
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z

; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6120988-A/No. 6120988-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0183.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1693 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-478-507-7

Query Match 32.8%; Score 45; DB 3; Length 1693;
Best Local Similarity 38.1%; Pred. No. 2.2e+02;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 SFCGAVWTLNRAIGHFVRGS 24
DB 647 AFCSALYRFNREARHSLIGN 667

RESULT 9
US-09-128-275A-7
; Sequence 7, Application US/09128275A
; Patent No. 6229005
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W

```

; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: NO. 6229005-A/No. 6229005-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,275A
; FILING DATE: 03-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joane R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1693 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-128-275A-7

```

Query Match 32.8%; Score 45; DB 3; Length 1693;
 Best Local Similarity 38.1%; Pred. No. 2.2e+02;
 Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```

QY 4 SFCGAVWTLNRAIGRHFRVGS 24
Db 647 AFCSALYRFNREARHSLIGN 667

```

RESULT 10
 US-09-553-427-7
 ; Sequence 7, Application US/09553427
 ; Patent No. 6379891
 ; GENERAL INFORMATION:
 ; APPLICANT: Reyes, Gregory R
 ; APPLICANT: Yarbough, Patrice O

```

; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: NO. 6379891-A/No. 6379891-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/553,427
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/478,507
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0183.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1693 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-553-427-7

```

Query Match 32.8%; Score 45; DB 3; Length 1693;
 Best Local Similarity 38.1%; Pred. No. 2.2e+02;
 Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```

QY 4 SFCGAVWTLNRAIGRHFRVGS 24
Db 647 AFCSALYRFNREARHSLIGN 667

```

RESULT 11
 US-09-673-395A-392
 ; Sequence 392, Application US/09673395A
 ; Patent No. 6620923

Search completed: November 10, 2004, 14:55:41
Job time : 10.2009 secs

INFORMATION FOR SEQ ID NO: 442:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-442

Query Match 32.1%; Score 44; DB 3; Length 63;
Best Local Similarity 58.3%; Pred. No. 7.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 VMTLNRAIGHF 20
DB 42 VMTLNNSCTHY 53

RESULT 15

US-09-201-945-442
Sequence 442, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 442:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-442

Query Match 32.1%; Score 44; DB 3; Length 63;
Best Local Similarity 58.3%; Pred. No. 7.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 VMTLNRAIGHF 20
DB 42 VMTLNNSCTHY 53

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 28.9151 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-57

Perfect score: 137

Sequence: 1 RPVSFCGAVTLNRAIGRHFVRGSR 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137	100.0	25	14	US-10-092-750-57
2	56.5	41.2	231	17	US-10-425-115-234613
3	53	38.7	168	16	US-10-437-963-163106
4	51	37.2	3042	16	US-10-437-963-123736
5	50.5	36.9	932	14	US-10-444-575-4
6	50	36.5	239	17	US-10-425-115-270985
7	49	35.8	1707	16	US-10-239-090A-12
8	49	35.8	1707	16	US-10-239-090A-54
9	49	35.8	1707	16	US-10-239-090A-56
10	47.5	34.7	91	17	US-10-425-115-256918
11	47.5	34.7	214	17	US-10-425-115-302062
12	47.5	34.7	536	17	US-10-425-115-187602
13	47	34.3	107	15	US-10-424-599-246506

14	47	34.3	235	15	US-10-424-599-246502	Sequence 246502,
15	47	34.3	247	15	US-10-425-114-40659	Sequence 40659, A
16	47	34.3	312	15	US-10-425-114-49039	Sequence 49039, A
17	47	34.3	392	15	US-10-425-114-43621	Sequence 43621, A
18	47	34.3	493	15	US-10-424-599-246507	Sequence 246507,
19	47	34.3	727	15	US-10-425-114-52764	Sequence 52764, A
20	47	34.3	727	15	US-10-425-114-55408	Sequence 55408, A
21	46	33.6	104	11	US-10-424-599-235252	Sequence 235252,
22	46	33.6	106	11	US-09-864-408A-4608	Sequence 4608, Ap
23	46	33.6	117	17	US-10-425-115-251991	Sequence 251991,
24	46	33.6	348	15	US-10-424-599-159122	Sequence 159122,
25	46	33.6	348	15	US-10-425-114-37391	Sequence 37391, A
26	46	33.6	394	14	US-10-369-493-12891	Sequence 12891, A
27	46	33.6	469	14	US-10-369-493-22836	Sequence 22836, A
28	46	33.6	536	16	US-10-437-963-162720	Sequence 162720,
29	46	33.6	551	14	US-10-223-277-3	Sequence 3, Appli
30	45.5	33.2	200	17	US-10-425-115-302416	Sequence 302416,
31	45	32.8	116	16	US-10-767-701-50399	Sequence 50399, A
32	45	32.8	161	14	US-10-029-386-33432	Sequence 33432, A
33	45	32.8	212	15	US-10-424-599-272781	Sequence 272781,
34	45	32.8	269	15	US-10-424-599-234166	Sequence 234166,
35	45	32.8	343	14	US-10-369-493-1046	Sequence 1046, Ap
36	45	32.8	438	14	US-10-156-761-13970	Sequence 13970, A
37	45	32.8	478	15	US-10-320-797-3206	Sequence 3206, Ap
38	45	32.8	541	9	US-09-815-242-10746	Sequence 10746, A
39	45	32.8	607	10	US-09-308-207-59	Sequence 59, Appli
40	45	32.8	1693	10	US-09-851-410-7	Sequence 7, Appli
41	44.5	32.5	69	16	US-10-437-963-160071	Sequence 160071,
42	44.5	32.5	86	17	US-10-425-115-196350	Sequence 196350,
43	44.5	32.5	97	17	US-10-425-115-264220	Sequence 264220,
44	44.5	32.5	160	15	US-10-424-599-238559	Sequence 238559,
45	44.5	32.5	463	15	US-10-389-586-1198	Sequence 1198, Ap

ALIGNMENTS

RESULT 1
US-10-092-750-57
; Sequence 57, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-57

Query Match 100.0%; Score 137; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.1e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVSFCGAVTLNRAIGRHFVRGSR 25

Db 1 RPVSFCGAVTLNRAIGRHFVRGSR 25

RESULT 2

US-10-425-115-234613

; Sequence 234613, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:


```
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(91)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_165898C.1.pep
US-10-425-115-256918

Query Match      34.7%; Score 47.5; DB 17; Length 91;
Best Local Similarity 57.9%; Pred. No. 18;
Matches 11; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 RVPVFCGAVW-TLNRAIGR 18
Db 17 RLVPVCGAVWLVLGGAVGR 35

RESULT 11
US-10-425-115-302062
; Sequence 302062, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 302062
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_38556C.1.pep
US-10-425-115-302062

Query Match      34.7%; Score 47.5; DB 17; Length 214;
Best Local Similarity 44.0%; Pred. No. 44;
Matches 11; Conservative 4; Mismatches 7; Indels 3; Gaps 1;

QY 1 RVPVFCGAVW-TLNRAIGR 22
Db 147 RPAHCSSCGSVFTSCQALGGHMR 171

RESULT 12
US-10-425-115-187602
; Sequence 187602, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 187602
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(536)

; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_102679C.1.pep
US-10-425-115-187602

Query Match      34.7%; Score 47.5; DB 17; Length 536;
Best Local Similarity 37.0%; Pred. No. 1,2e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 5; Gaps 1;

QY 2 PVFCGAVWTLNRAIGRHF-----VRG 23
Db 240 PIMFCSLKTQNRXSRHFFHLHCIRG 266

RESULT 13
US-10-424-599-246506
; Sequence 246506, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 246506
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64626C.1.pep
US-10-424-599-246506

Query Match      34.3%; Score 47; DB 15; Length 107;
Best Local Similarity 57.1%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 GAVWTLNRAIGRHF 20
Db 26 GEVWNGRSIGRHW 39

RESULT 14
US-10-424-599-246502
; Sequence 246502, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 246502
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(235)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64622C.1.pep
US-10-424-599-246502
```

```

Query Match      34.3%; Score 47; DB 15; Length 235;
Best Local Similarity 57.1%; Pred. No. 58;
Matches      8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      7 GAVWTLNRAIGRHF 20
Db      154 GEVWNGRSIGRHW 167

RESULT 15
US-10-425-114-40659
; Sequence 40659, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313) B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40659
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700846835_FLI.pep
US-10-425-114-40659

Query Match      34.3%; Score 47; DB 15; Length 247;
Best Local Similarity 57.1%; Pred. No. 61;
Matches      8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      7 GAVWTLNRAIGRHF 20
Db      166 GEVWNGRSIGRHW 179

Search completed: November 11, 2004, 02:43:13
Job time : 29.9651 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 6.13208 Seconds
(without alignments)

Title: US-10-092-750-57

Perfect score: 137
Sequence: 1 RPVSFCGAVWTLNRAIGRHFVGRSR 25

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *

```
1: pir1:*
```

```
2: pir2:*
```

```
3: pir3:*
```

```
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	51	37.2	37.6	2	S42746	2	fibrial adhesin f
2	50.5	36.9	932	2	JCS953	3	inter-alpha-inhibi
3	50	35.8	496	2	T15691	1	hypothetical prote
4	49	35.5	376	2	S42747	1	fibrial adhesin f
5	43	35.8	376	2	S36247	1	fibrial adhesin f
6	48	35.0	195	2	T32807	1	hypothetical prote
7	48	35.0	366	2	T54401	1	hypothetical prote
8	48	35.0	407	2	D87534	4	hypothetical prote
9	48	35.0	557	2	T41588	3	probable amino-ac
10	47	34.3	318	2	A86750	1	hypothetical prote
11	46	33.6	237	2	AH1855	3	hypothetical prote
12	45	33.6	237	2	S14722	3	hypothetical prote
13	45	33.6	273	2	S69193	3	probable finger pr
14	45	33.6	455	2	P83460	1	probable amino aci
15	45	33.6	459	2	D72722	2	probable MRSA prot
16	45	33.6	472	2	AH0109	3	probable histidine
17	45.5	33.2	362	2	C82070	2	conserved hypothet
18	45	32.8	256	2	E70868	2	probable enoyl-coA
19	45	32.8	315	2	S44019	3	aerobactin biosynt
20	45	32.8	315	2	T44480	3	conserved hypothet
21	45	32.8	340	2	E81331	1	uroporphyrinogen d
22	45	32.8	343	2	A64443	3	glyceralddehyde-3-p
23	45	32.8	417	2	A81919	1	Probable sodium-tr
24	45	32.8	529	2	JE0181	1	oligo-1,6-glucosid
25	45	32.8	668	2	AH1565	1	hypothetical prote
26	45	32.8	1693	1	MNW9HE	1	genome polyprotein
27	44.5	32.5	351	2	A88103	3	protein W10G11.7 [
28	44.5	32.5	463	2	E95377	1	probable OtsA treh
29	44.5	32.5	1197	2	D82696	1	hypothetical prote

ALIGNMENTS

RESULT 1

R50011
 S42746
 fibrial adhesin fimD - Bordetella parapertussis
 C\$Species: Bordetella parapertussis
 C\$Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C\$Accession: S42746
 R\$Willems, R.J.L.
 submitted to the EMBL Data Library, October 1993
 A\$Reference number: S42746
 A\$Accession: S42746
 A\$Status: preliminary
 A\$Molecule type: DNA
 A\$Residues: 1-376 <WIL>

```

Query Match      37.2%; Score 51; DB 2; Length 376;
Best Local Similarity 36.8%; Pred. NO. 4.4;
Matches 14; Conservative 1; Mismatches 5; Indels 18; Gaps 2;

QY      1  RPVSFCG---AWWT-----LNRAIGHRF 20
          :|:|:|:|:|:|:|:|:|:|:|:|
Db      15  RVISFGGAALVWAGLAVOPAMAVDPVDCGRAIHLHF 52
          :|:|:|:|:|:|:|:|:|:|:|:|

```

RESULTS

inter-alpha-inhibitor H4p heavy chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JCS953
R;Soury, E.; Olivier, E.; Daveau, M.; Hiron, M.; Claeysens, S.; Risler, J.L.; Salier, Biochem. Biophys. Res. Commun. 243, 522-530, 1998
A>Title: The H4p heavy chain of inter-alpha-inhibitor family largely differs in the structure of the inter-alpha-inhibitor complex component II
A:Reference number: JCS953; MUID:98153798; PMID:9480842
A:Accession: JCS953
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-932 <SOU>
A:Cross-references: UNIPROT:O35802; GB:Y11283; NID:G2292987; PIDN:CAA72155.1; PID:G2292987
C:Superfamily: inter-alpha-trypsin inhibitor complex component II

Query Match	36.9%	Score 50.5;	DB 2;	Length 932;
Best Local Similarity	28.9%	Pred. No. 13;		
Matches	11.	Conservative	6;	Mismatches
			8;	Indels
			13;	Gaps

QY 1 RPVSFCGAVWTLNRAIGRHF-----VRGSR 25
: | | | | : | | : | | : | |
606 KPTVEVGGCVWSTLSAVORPHKPTPTTGSKILLTSRLRGR 643

FROM E

RESULT
T75691

hypothetical protein C28G1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T15691
R;Favell, T.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C28G1.
A;Reference number: Z18389
A;Accession: T15691
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-496 <FAV>
A;Cross-references: UNIPROT:Q18288; EMBL:U41026; NID:g1086701; PID:g1086704; PIDN:AAA823
C;Genetics:
A;Gene: CBSP:C28G1.1
A;Introns: 75/3; 102/3; 206/2; 382/3; 414/3
Query Match 36.5%; Score 50; DB 2; Length 496;
Best Local Similarity 58.8%; Pred. No. 8.2;
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 8 AVWTLNRAIGRHVGRGS 24
Db 4 ADWNLERAIGTHLAGS 20
RESULT 4
S42747
fimbrial adhesin fimD - Bordetella bronchiseptica
C;Species: Bordetella bronchiseptica
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S42747
R;Willems, R.J.L.
submitted to the EMBL Data Library, October 1993
A;Reference number: S42746
A;Accession: S42747
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-376 <WIL>
A;Cross-references: UNIPROT:Q44887; EMBL:X75811; NID:g416478; PIDN:CAA53446.1; PID:g4164
Query Match 35.8%; Score 49; DB 2; Length 376;
Best Local Similarity 34.2%; Pred. No. 8.9;
Matches 13; Conservative 2; Mismatches 5; Indels 18; Gaps 2;
QY 1 RPVSFCG----AVWT-----LNRAIGRHF 20
Db 15 RVISFCGAALAVWAGLVQPMAMVDPVDCGRALGLHF 52
RESULT 5
S36247
fimbrial adhesin fimD precursor - Bordetella pertussis
N;Alternate names: filamentous hemagglutinin fhae
C;Species: Bordetella pertussis
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S36247; S61765; S22871; S23565
R;Willems, R.J.L.; Geuijen, C.; van der Heide, H.G.J.; Macheson, M.; Robinson, A.; Versl
Mol. Microbiol. 9, 623-634, 1993
A;Title: Isolation of a putative fimbrial adhesin from Bordetella pertussis and the iden
A;Reference number: S36244; MUID:94018656; PMID:8105363
A;Accession: S36247
A;Molecule type: DNA
A;Residues: 1-376 <WIL>
A;Cross-references: UNIPROT:Q00879; EMBL:X64876; NID:g313839; PIDN:CAA46091.1; PID:g3974
A;Accession: S61765
A;Molecule type: protein
A;Residues: 289-308;332-363 <LOW>
R;Locht, C.; Geoffroy, M.; Renauld, G.
submitted to the EMBL Data Library, June 1992
A;Description: Common accessory genes for the Bordetella pertussis filamentous hemagglut
A;Reference number: S22869
A;Status: preliminary

A;Accession: S22871
A;Molecule type: DNA
A;Residues: 1-376 <LOC>
A;Cross-references: EMBL:X66729; NID:g39716; PIDN:CAA47267.1; PID:g39719
A;Genetics: FHA
R;Locht, C.; Geoffroy, M.C.; Renauld, G.
EMBO J. 11, 3175-3183, 1992
A;Title: Common accessory genes for the Bordetella pertussis filamentous hemagglutinin
A;Reference number: S23563; MUID:92371423; PMID:1354611
A;Accession: S23565
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 216-376 <LOF>
A;Cross-references: EMBL:X66729
A;Genetics: FHA
C;Genetics: <FIM>
A;Gene: fimD
C;Genetics: <FHA>
A;Gene: fhaB
C;Keywords: blocked amino end
F;1-37/Domain: signal sequence #status predicted <SIG>
F;38-376/Product: fimbrial adhesin fimD #status experimental <XAT>
Query Match 35.8%; Score 49; DB 2; Length 376;
Best Local Similarity 34.2%; Pred. No. 8.9;
Matches 13; Conservative 2; Mismatches 5; Indels 18; Gaps 2;
QY 1 RPVSFCG----AVWT-----LNRAIGRHF 20
Db 15 RVISFCGAALAVWAGLVQPMAMVDPVDCGRALGLHF 52
RESULT 6
T32807
hypothetical protein F52C6.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32807
R;Wamsley, P.; Gibson, A.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid F52C6.
A;Reference number: Z21226
A;Accession: T32807
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-195 <WAM>
A;Cross-references: UNIPROT:O44821; EMBL:AF040645; PIDN:AAB94971.1; GSPDB:GN000020; CESP
A;Experimental source: strain Bristol N2; clone F52C6
C;Genetics:
A;Gene: CBSP:F52C6.2
A;Map position: 2
A;Introns: 38/1; 110/2
Query Match 35.0%; Score 48; DB 2; Length 195;
Best Local Similarity 58.3%; Pred. No. 6.6;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 3 VSFCGAVWTINR 14
Db 178 LSYCGALWTSR 189
RESULT 7
T51401
hypothetical protein F14F8.100 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51401
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asanizu, E.; Kotani, H.; Tabata, S.; Me
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51401
A;Status: preliminary

A:Molecule type: DNA
A:Residues: 1-366 <SAT>
A:Cross-references: UNIPROT:Q9LFV4; EMBL:AL391144
A:Experimental source: cultivar Columbia; BAC clone F14F8
C:Genetics:
A:Map position: 5
A:Introns: 80/1; 124/3; 207/3; 294/1
A>Note: F14F8_100
C:Superfamily: myrosinase-associated protein MyAP

Query Match 35.0%; Score 48; DB 2; Length 366;
Best Local Similarity 40.0%; Pred. No. 12;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
QY 1 RVSPFCGAVWTLNRAIGHFVRGSR 25
DB 103 RGVNYASAAAGILDETGRHYRGAR 127

RESULT 8

D87534
Hypothetical protein CC2301 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: D87534

R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.B.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: D87534

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-407 <STO>

A:Cross-references: UNIPROT:Q9A5Z6; GB:AE005673; NID:gi3423820; PIDN:AAK24272.1; GSPDB:G

C:Genetics:

A:Gene: CC2301

Query Match 35.0%; Score 48; DB 2; Length 407;
Best Local Similarity 52.9%; Pred. No. 14;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 7 GAVWTLNRAIGHFVRG 23

DB 229 GVAWLGTRVICKPFVRG 245

RESULT 9

T41588
Probable amino-acid permease - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T41588

R.Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z21968

A:Accession: T41588

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-557 <MUR>

A:Cross-references: UNIPROT:Q74537; EMBL:AL031543; PIDN:CAA20834.1; GSPDB:GN00068; SPDB:

A:Experimental source: strain 972h-; cosmid c74

C:Genetics:

A:Gene: SPDB:SPCC74.04

A:Map position: 3

C:Superfamily: choline transport protein

Query Match 35.0%; Score 48; DB 2; Length 557;
Best Local Similarity 54.5%; Pred. No. 19;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 VSPFCGAVWTLN 13

DB 265 MSFCGVITMS 275

RESULT 10

A86750
Hypothetical protein ykcG [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86750

R;Bolojin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarma, K.; Weissenbach, J.; Ehrli

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: A86750

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-318 <STO>

A:Cross-references: UNIPROT:Q9CGU1; GB:AE005176; PID:gl2723948; PIDN:AAK05099.1; GSPDB:

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: ykcG

C:Superfamily: stress response protein csbB

Query Match 34.3%; Score 47; DB 2; Length 318;
Best Local Similarity 33.3%; Pred. No. 15;
Matches 10; Conservative 5; Mismatches 9; Indels 6; Gaps 1;

QY 1 RVPSPCGAVWTLNRAIG-----RHPVRGS 24

DB 237 RLFWIGLIWFLIAIVGAIVIAIARHPINGS 266

RESULT 11

AH1855

Hypothetical protein all0393 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AH1855

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH1855

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-193 <KUR>

A:Cross-references: UNIPROT:Q8YZR4; GB:BA000019; PIDN:BAE72351.1; PID:gl17129738; GSPDB:

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all0393

Query Match 33.6%; Score 46; DB 2; Length 193;
Best Local Similarity 31.8%; Pred. No. 13;
Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 2 PVSFCGAVWTLNRAIGHF 20

DB 81 PMSYWSYWTIDIVARHY 99

RESULT 12

S14722

Hypothetical protein - bovine

C:Species: Bos primigenius indicus x Bos primigenius taurus (cattle)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Jul-1999

C:Accession: S14722

R;Tanaka, M.; Minoura, H.; Ushiro, H.; Nakashima, K.

Biochim. Biophys. Acta 1088, 385-389, 1991

A:Title: A novel cDNA clone encoding a prolactin-like protein that lacks the two C-term

A:Reference number: S14722; MUID:91198142; PMID:2015300

A;Accession: S14722

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-237 <FAN>

A;Cross-references: GB:X59504; NID:G561; PIDN:CAA42092.1; PID:G562

C;Superfamily: prolactin

Query Match 33.6%; Score 46; DB 2; Length 237;

Best Local Similarity 52.4%; Pred. No. 16;

Matches 11; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 4 SFGCAVWTLNRAIGRHFVRGS 24

||||| |||

Db 6 SFGCHQWYNTNP-----VRGS 20

RESULT 13

S69193

probable finger protein Pszf1 - garden pea

C;Species: Pisum sativum (garden pea)

C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C;Accession: S69193

R;Michael, A.J.; Hofer, J.M.I.; Ellis, T.H.N.

Plant Mol. Biol. 30, 1051-1058, 1996

A;Title: Isolation by PCR of a cDNA clone from pea petals with similarity to petunia and

A;Reference number: S69193; MUID:96270382; PMID:8639742

A;Accession: S69193

A;Molecule type: mRNA

A;Residues: 1-273 <MIC>

A;Cross-references: UNIPROT:Q41070; EMBL:X87374; NID:G854399; PIDN:CAA60828.1; PID:G8610

C;Superfamily: Arabidopsis thaliana hypothetical protein F1ZE4.290

C;Keywords: DNA binding; zinc finger

F;115-135/Region: zinc finger CCHH motif

F;202-222/Region: zinc finger CCHH motif

Query Match 33.6%; Score 46; DB 2; Length 273;

Best Local Similarity 47.4%; Pred. No. 19;

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 SFGCAVWTLNRAIGRHFVR 22

||||| :|||

Db 203 SICGAEFTSQALGGHMR 221

RESULT 14

F83460

probable amino acid permease PA1485 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: F83460

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: F83460

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-455 <STO>

A;Cross-references: UNIPROT:Q913M7; GB:AE004577; GB:AE004091; NID:G9947430; PIDN:AAG0487

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA1485

Query Match 33.6%; Score 46; DB 2; Length 455;

Best Local Similarity 52.6%; Pred. No. 31;

Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 PVSFCGAVWTLNRAIGRHF 20

||||| |||

Db 109 PVLATGAATYLNALGTHF 127

RESULT 15

D72722

probable MRSA protein APE0317 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: D72722

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-DO, K.; Takai

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; I

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: D72722

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-469 <KAW>

A;Cross-references: UNIPROT:Q9YFC4; DBJ:AF000059; NID:G5103911; PIDN:BAA79272.1; PID:G

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE0317

C;Superfamily: phosphomannomutase

Query Match 33.6%; Score 46; DB 2; Length 469;

Best Local Similarity 64.7%; Pred. No. 32;

Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 9 VWTINRAIGRHFVRGSR 25

||||| :|||

Db 30 VLRLARAIGAYFGRGSR 46

Search completed: November 10, 2004, 14:52:25

Job time : 7.13208 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 33.3019 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-57

Perfect score: 137
Sequence: 1 RPVSFCGAVTINRAIGHRHFVRGSR 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt.02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	38.7	84	2	Q8KB37
2	52	38.0	791	2	Q8Z223
3	52	38.0	1693	2	Q9WLL5
4	52	38.0	2503	2	Q8B912
5	51	37.2	365	2	Q7W695
6	51	37.2	376	2	Q44888
7	51	37.2	1693	2	Q69418
8	50.5	36.9	932	2	Q35802
9	50	36.5	199	2	Q7Q1J4
10	50	36.5	496	2	Q18288
11	50	36.5	1706	2	Q7TGF2
12	49.5	36.1	573	2	Q9V516
13	49	35.8	341	2	P93714
14	49	35.8	365	2	Q7W164
15	49	35.8	376	1	FAAE_BORPE
16	49	35.8	376	2	Q44887
17	49	35.8	389	2	Q6GXX0
18	49	35.8	389	2	Q6GXX1
19	49	35.8	389	2	Q6GXX2
20	49	35.8	389	2	Q6GXX3
21	49	35.8	1684	2	Q8JUM2
22	49	35.8	1707	2	Q6PMR4
23	49	35.8	1707	2	Q7TA93
24	49	35.8	1707	2	Q806D8
25	49	35.8	1707	2	Q806E1
26	49	35.8	1707	2	Q801R6
27	49	35.8	1707	2	Q8JUN3
28	49	35.8	1707	2	Q8JUN7
29	49	35.8	1707	2	Q9IVZ9
30	49	35.8	1707	2	AA02424
31	48	35.0	90	2	Q9WD34

32 48 35.0 165 2 Q8GIM9 uncultured
33 48 35.0 366 2 Q9LFV4 arabidopsis
34 48 35.0 407 2 Q9ASZ6 caulobacter
35 48 35.0 492 2 Q8FUF9 bruceella su
36 48 35.0 512 2 Q9JH69 turkey astr
37 48 35.0 523 2 Q8WOH1 oryza sativ
38 48 35.0 557 1 YCQ4 SCHPO
39 48 35.0 1199 2 Q7PS03 anopheles g
40 47 34.3 152 2 Q84EA5 uncultured
41 47 34.3 223 2 Q88Y79 lactobacill
42 47 34.3 262 2 Q89K29 bradyrhizob
43 47 34.3 318 2 Q9CGU1 lactococcus
44 47 34.3 424 2 Q7UWK8 rhodospirill
45 47 34.3 452 2 Q82SE3 nitrosomona

ALIGNMENTS

RESULT 1

Q8KB37 PRELIMINARY; PRT; 84 AA.
ID Q8KB37
AC Q8KB37
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein CT1960.
GN OrderedLocusNames=CT1960;
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobaculum.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.P., Wu M.,
RA Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
RA Radune D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,
RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
EMBL; AE012946; AAM73179.1; -
TIGR; CT1960; -
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; Aldo ket red; 1.
DR ProDom; PD000288; Aldo/Ket red; 1.
DR Complete Proteome; Hypothetical protein.
KW Complete Proteome; Hypothetical protein.
SQ SEQUENCE 84 AA; 9282 MW; 7246A595F4A9CE56 CRC64;

Query Match 38.7%; Score 53; DB 2; Length 84;

Best Local Similarity 57.9%; Pred. No. 3.5;

Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

Qy 1 RPVSF--CGAVTINRAIG 17

Db 33 KPLVFTKCGLVWDENRAIG 51

RESULT 2

Q8Z223 PRELIMINARY; PRT; 791 AA.
ID Q8Z223
AC Q8Z223
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein PA30478.
GN OrderedLocusNames=PA30478;
OS Pyrobaculum aerophilum.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-587;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi1227;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cardeno-Taraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moul S., Norbercaak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640432; CAE39313.1; -;
 DR GO; GO:0009289; C:fimbria; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR000259; Fimbrial.
 DR Pfam; PF00419; Fimbrial; 1.
 KW Complete proteome.
 SQ SEQUENCE 365 AA; 39057 MW; 0726DF5B2295D6B CRC64;

Query Match 37.2%; Score 51; DB 2; Length 365;
 Best Local Similarity 36.8%; Pred. No. 30;
 Matches 14; Conservative 1; Mismatches 5; Indels 18; Gaps 2;

OY 1 RPVSFCG---AVWT-----LNRAIGRHF 20
 DB 4 RVISFGAALAVAGLAVQPMADVDPVDCGRAIGLHF 41

RESULT 6
 ID Q44888 PRELIMINARY; PRT; 376 AA.
 AC Q44888;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE FimD fimbrial adhesin (Put.) precursor.
 GN Name=fimD;
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BPP24;
 RX MEDLINE=94018656; PubMed=8105363;
 RA Willems R.J.L., Geuljen C., van der Heide H.G.J., Matheson M.,
 RA Robinson A., Versluis L.F., Eberink R., Theelen J., Mooi F.R.;
 RT "Isolation of a putative fimbrial adhesin from Bordetella pertussis
 RT and the identification of its gene";
 RL Mol. Microbiol. 9:623-634(1993).
 DR EMBL; X75812; CAA53447.1; -;
 DR PIR; S42746; S42746;
 DR GO; GO:0009289; C:fimbria; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR008966; Adhes.bact.
 DR InterPro; IPR000259; Fimbrial.
 DR Pfam; PF00419; Fimbrial; 1.
 KW SIGNAL.
 FT SIGNAL 1 37 Potential.
 FT CHAIN 38 376 FimD fimbrial adhesin (put.).
 SQ SEQUENCE 376 AA; 40297 MW; BAC5COAB447C9F06 CRC64;

Query Match 37.2%; Score 51; DB 2; Length 376;
 Best Local Similarity 36.8%; Pred. No. 30;

Matches 14; Conservative 1; Mismatches 5; Indels 18; Gaps 2;

OY 1 RPVSFCG---AVWT-----LNRAIGRHF 20
 DB 4 RVISFGAALAVAGLAVQPMADVDPVDCGRAIGLHF 41

RESULT 7
 ID Q69418 PRELIMINARY; PRT; 1693 AA.
 AC Q69418;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE ORF1, ORF2 & ORF3.
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OC NCBI_TaxID=12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA von Brunn A., Seebach J., Thyagarajan S.P., Mohanavalli B., Menon T.,
 RA Friesner G.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X99441; CAA67802.1; -;
 DR GO; GO:0008174; F:RNA methyltransferase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003724; F:RNA helicase activity; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0016032; P:viral life cycle; IEA.
 DR GO; GO:0019082; P:viral protein processing; IEA.
 DR InterPro; IPR002589; Alpp.
 DR InterPro; IPR008748; Peptidase_C41.
 DR InterPro; IPR001788; RNA_dep_RNAPol2.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PStir.
 DR InterPro; IPR006066; Viral_helicase.
 DR InterPro; IPR002588; V_methyltrans.
 DR Pfam; PF01661; Alpp; 1.
 DR Pfam; PF05417; Peptidase_C41; 1.
 DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
 DR Pfam; PF01443; Viral_helicase1; 1.
 DR Pfam; PF01660; Vmethyltransf; 1.
 DR SMART; SM00506; Alpp; 1.
 SQ SEQUENCE 1693 AA; 185949 MW; 489B4ADBA5E7E529 CRC64;

Query Match 37.2%; Score 51; DB 2; Length 1693;
 Best Local Similarity 38.1%; Pred. No. 1.3e+02;
 Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 1 RPVSFCG---AVWT-----LNRAIGRHF 20
 DB 15 RVISFGAALAVAGLAVQPMADVDPVDCGRAIGLHF 52

RESULT 8
 ID Q35802 PRELIMINARY; PRT; 932 AA.
 AC Q35802;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Inter-alpha-inhibitor H4 heavy chain.
 GN Name=ITIH4;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Non-structural proteins.
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22811791; PubMed=12931035;
 RA Liu Z., Chi B., Takahashi K., Mishihiro S.;
 RT "A genotype IV hepatitis E virus strain that may be indigenous to
 ET Changchun, China.";
 RL Intervirology 46:252-255(2003).
 DR EMBL; AB108537; BAC77167.1; -;
 DR GO; GO:0008174; F:RNA binding; IEA.
 DR GO; GO:0003723; F:RNA helicase activity; IEA.
 DR GO; GO:0003724; F:RNA helicase activity; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0016032; P:viral life cycle; IEA.
 DR GO; GO:0019082; P:viral protein processing; IEA.
 DR InterPro; IPR002589; Alpp.
 DR InterPro; IPR008748; Peptidase C41.
 DR InterPro; IPR001788; RNA_dep_RNApol2.
 DR InterPro; IPR007095; RNA_pol_DS_Ps.
 DR InterPro; IPR007094; RNA_pol_PsVir.
 DR InterPro; IPR006066; Viral helicase1.
 DR InterPro; IPR002588; V methyltrans.
 DR Pfam; PF01661; Alpp; 1.
 DR Pfam; PF05417; Peptidase C41; 1.
 DR Pfam; PF00978; RNA_dep_RNApol2; 1.
 DR Pfam; PF01443; Viral helicase1; 1.
 DR Pfam; PF01660; Vmethyltransf; 1.
 DR SMART; SM00506; Alpp; 1.
 SQ SEQUENCE 1706 AA; 187032 MW; 4980167080EAD21C CRC64;

 Query Match 36.5%; Score 50; DB 2; Length 1706;
 Best Local Similarity 45.0%; Fred. No. 1.9e+02;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

 QY 4 SFGGAVTFLNRAIGRHFFVRG 23
 Db 647 AFCSALYFNRICQHSILIG 666

 RESULT 12
 Q9V516 ID Q9V516 PRELIMINARY; PRT; 573 AA.
 AC Q9V516
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE CG8172-PA.
 GN ORFNames=CG8172;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426089; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tu P.J., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG FLYBASE;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FLYBASE;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to peptidase family S1.

```

DE EMBL; AE003835; AAF59006.1; -.
GN HSP; P00750; IRTF.
OS FlyBase; FBgn0033362; CG8172.
RA GO; GO:0004295; P:trypsin activity; NAS.
RA GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYSP; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 573 AA; 62084 MW; BCIDA36CD97BEE0 CRC64;

Query Match 36.1%; Score 49.5; DB 2; Length 573;
Best Local Similarity 55.0%; Pred. No. 77;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 RPVSFCGAVWT-LNRAIGH 19
DB 285 RVPVCGEVYTRSRIVGCH 304

RESULT 13
P93714
ID P93714 PRELIMINARY; PRT; 341 AA.
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PETH; ZPT2-6.
OS Petunia hybrida (Petunia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Petunia.
OX NCBI_TaxID=4102;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stamen.
RA Kobayashi A., Sakamoto A., Kubo K., Rybka Z., Kanno Y., Takatsuji H.;
RA Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB000452; BAA19111.1; -.
DR HSP; Q38895; INJQ.
DR TRANSFAC; T03943; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; ZnF_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR SMART; SM00355; ZnF_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
SQ SEQUENCE 341 AA; 38431 MW; 4122B0C61505D1F4 CRC64;

Query Match 35.8%; Score 49; DB 2; Length 341;
Best Local Similarity 47.4%; Pred. No. 56;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 SFCGAVTTLNRAIGHFVR 22
DB 243 SYCGAFTSGQALGHMRR 261

RESULT 14
Q7WI64
ID Q7WI64 PRELIMINARY; PRT; 365 AA.
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Fimbrial adhesin.
GN Name=fimD; Synonyms=fhaE; OrderedLocusNames=BB2989;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RC MEDLINE=2287954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Shimmonds M., Skelton J., Squares R., Stevens K.,
RA Unwin L., Whitehead S., Barrrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640446; CAE33481.1; -.
DR GO; GO:0009289; C:fimbria; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome.
SQ SEQUENCE 365 AA; 39069 MW; 25199A15E2127A2C CRC64;

Query Match 35.8%; Score 49; DB 2; Length 365;
Best Local Similarity 34.2%; Pred. No. 59;
Matches 13; Conservative 2; Mismatches 5; Indels 18; Gaps 2;

QY 1 RPVSFCG---AVWT-----LNRAIGHF 20
DB 4 RVISFGAALVWAGLVQFAMVDPVDCGRLGHF 41

RESULT 15
FHA_E BORPE STANDARD; PRT; 376 AA.
ID PHAE BORPE
AC Q00879;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Protein fhaE precursor.
GN Name=fhaE; Synonyms=fimD; OrderedLocusNames=BP1883;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RC MEDLINE=92371423; PubMed=1354611;
RA Loch C., Geoffroy M.C., Renaud G.;
RA "Common accessory genes for the Bordetella pertussis filamentous
RT hemagglutinin and fimbriae share sequence similarities with the papC
RT and papD gene families.";
RL EMBO J. 11:3175-3183(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=94018656; PubMed=8105363;
RA Willens R.J.L., Geuijen C., van der Heide H.G.J., Matheson M.,
RA Robinson A., Verschuif L.F., Ebberink R., Theelen J., Mooi F.R.;
RA "Isolation of a putative fimbrial adhesin from Bordetella pertussis
RT and the identification of its gene.";
RL Mol. Microbiol. 9:623-634(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;

```



```

RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X66729; CAA47267.1; -.
DR EMBL; X64876; CAA46031.1; -.
DR EMBL; BX640416; CAA42166.1; ALT_INIT.
DR PIR; S36247; S36247.
DR InterPro; IPR008966; Adhes.bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
KW Complete proteome; Signal.
FT SIGNAL 1 37 Potential.
FT CHAIN 38 376 Protein fhaE.
SQ SEQUENCE 376 AA; 40528 MW; 2820D28FCBDB4ECF CRC64;
Query Match 35.8%; Score 49; DB 1; Length 376;
Best Local Similarity 34.2%; Pred. No. 61;
Matches 13; Conservative 2; Mismatches 5; Indels 18; Gaps 2;
OY 1 RVSFSGC---AVWT-----LNRAIGRHF 20
Db 15 RVISFGAALAVAGLAVQPMADVDPVDCGRLGLHF 52

```

Search completed: November 10, 2004, 14:50:24
Job time : 35.3019 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 10.6151 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-58

PERFECT score: 144
Sequence: 1 HAVVARLLHIGAIMFORLDFIEOLSAPPA 29

Scoring table: BLOSUM62

Searched: 478139 seqs. 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

```

Database :
Issued Patents_AA.*
1: /cgn2_6/pdata/1/iaa/5A_COMB pep.*
2: /cgn2_6/pdata/1/iaa/5B_COMB pep.*
3: /cgn2_6/pdata/1/iaa/6A_COMB pep.*
4: /cgn2_6/pdata/1/iaa/6B_COMB pep.*
5: /cgn2_6/pdata/1/iaa/PCTUS_COMB.pe
6: /cgn2_6/pdata/1/iaa/backfiles1.pe

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	53	36.8	441	4	US-09-270-767-47360
2	46.5	32.3	184	4	US-10-140-002-76
3	46	31.9	106	4	US-09-270-767-58137
4	46	31.9	552	4	US-09-801-191A-2
5	46	31.9	552	4	US-09-801-191A-6
6	46	31.9	576	4	US-09-801-191A-4
7	46	31.9	576	4	US-09-801-191A-5
8	45.5	31.6	1233	4	US-09-328-352-7874
9	45	31.2	271	4	US-09-252-991A-44319
10	45	31.2	773	4	US-09-252-991A-31403
11	45	31.2	1132	2	US-08-567-508C-2
12	45	31.2	1132	3	US-09-196-480-2
13	45	31.2	1132	4	US-09-972-800A-16
14	45	31.2	1139	3	US-09-046-158A-22
15	44	30.6	87	4	US-09-543-681A-6899
16	44	30.6	147	4	US-09-248-796A-37890
17	44	30.6	220	3	US-08-297-431B-31
18	44	30.6	221	3	US-08-297-431B-2
19	44	30.6	221	3	US-08-297-431B-4
20	44	30.6	221	3	US-08-297-431B-6
21	44	30.6	221	3	US-08-297-431B-8
22	44	30.6	221	3	US-08-297-431B-10
23	44	30.6	221	3	US-08-297-431B-12
24	44	30.6	221	3	US-08-297-431B-14
25	44	30.6	221	3	US-08-297-431B-16
26	44	30.6	221	3	US-08-297-431B-18
27	44	30.6	221	3	US-08-297-431B-20

28	43.5	30.2	379	4	US-09-491-577-6	Sequence 6, Appl1
29	43.5	30.2	450	4	US-09-282-991A-31528	Sequence 31528, A
30	43	29.9	445	4	US-09-282-991A-26149	Sequence 26149, A
31	43	29.9	1722	4	US-09-538-092-1033	Sequence 1033, Ap
32	42.5	29.5	235	4	US-09-270-767-41614	Sequence 41614, A
33	42.5	29.5	1174	4	US-09-282-991A-29279	Sequence 29279, A
34	42	29.2	220	4	US-09-328-352-5959	Sequence 5959, Ap
35	42	29.2	285	4	US-09-252-991A-19996	Sequence 19996, A
36	42	29.2	385	4	US-09-252-991A-30436	Sequence 30436, A
37	42	29.2	355	3	US-09-134-001C-3580	Sequence 3580, Ap
38	42	29.2	510	4	US-09-707-767-46399	Sequence 46399, A
39	42	29.2	550	4	US-08-252-991A-17682	Sequence 17682, A
40	42	29.2	628	2	US-08-394-477-6	Sequence 6, Appl1
41	42	29.2	628	3	US-08-394-912A-6	Sequence 6, Appl1
42	42	29.2	628	4	US-09-333-636-6	Sequence 6, Appl1
43	42	29.2	711	4	US-09-252-991A-19682	Sequence 19682, A
44	42	29.2	774	4	US-09-252-991A-23814	Sequence 19984, A
45	42	29.2	836	4	US-09-282-991A-23513	Sequence 23513, A

ALIGNMENTS

```

RESULT 1
US-09-270-767-47360
; Sequence 47360, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47360
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47360

```

```
Query Match      36.8%; Score 53; DB 4; Length 441;
Best Local Similarity 47.4%; Pred. No. 2.2;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
```

Qy. 9 HIGAIMFQRLDFIEQLSAP 27
|| ::|: :| |||||
Db 37 HIESLLFRAIDTIEOLOKP 55

```

RESULT 2
US-10-140-002-76
/ Sequence 76, Application US/10140002
/ Patent No. 6725730
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Demoyere, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zewin
/ TITLE OF INVENTION: SECURED AND TR

```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RJC59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 76
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-002-76

Query Match 32.3%; Score 46.5; DB 4; Length 184;
Best Local Similarity 48.3%; Pred. No. 8.7;
Matches 14; Conservative 3; Mismatches 9; Indels 3; Gaps 1;

QY 4 VVARLLHIGAIMFQRL---LDFIEQLSAPPA 29
DB 93 VVALQHLKALDLRSNQFQDFPEQLTALPA 121

RESULT 3
US-09-270-767-58137
; Sequence 58137, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58137
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-58137

Query Match 31.9%; Score 46; DB 4; Length 106;
Best Local Similarity 45.0%; Pred. No. 5.4;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 HAVARLLHIGAIMFQRLDF 20
DB 87 HLTGQNLHENAATVTEKLD 106

RESULT 4
US-09-801-191A-2
; Sequence 2, Application US/09801191A
; Patent No. 6537788
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001159
; CURRENT APPLICATION NUMBER: US/09/801,191A
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Human
US-09-801-191A-2

Query Match 31.9%; Score 46; DB 4; Length 552;
Best Local Similarity 36.7%; Pred. No. 39;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY 3 VVARLLHIGAIMFQRL---DFIEQLSAPP 28
DB 187 VIARILHGMVAQQGLLHVGDIIKEVNGQP 192

RESULT 5
US-09-801-191A-6
; Sequence 6, Application US/09801191A
; Patent No. 6537788
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001159
; CURRENT APPLICATION NUMBER: US/09/801,191A
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Mus Musculus
US-09-801-191A-6

Query Match 31.9%; Score 46; DB 4; Length 552;
Best Local Similarity 36.7%; Pred. No. 39;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY 3 VVARLLHIGAIMFQRL---DFIEQLSAPP 28
DB 163 VIARILHGMVAQQGLLHVGDIIKEVNGQP 192

RESULT 6
US-09-801-191A-4
; Sequence 4, Application US/09801191A
; Patent No. 6537788
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001159
; CURRENT APPLICATION NUMBER: US/09/801,191A
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Human
US-09-801-191A-4

Query Match 31.9%; Score 46; DB 4; Length 576;
Best Local Similarity 36.7%; Pred. No. 42;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY 3 VVARLLHIGAIMFQRL---DFIEQLSAPP 28
DB 187 VIARILHGMVAQQGLLHVGDIIKEVNGQP 216

RESULT 7
US-09-801-191A-5
; Sequence 5, Application US/09801191A
; Patent No. 6537788
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001159
; CURRENT APPLICATION NUMBER: US/09/801,191A

```

```
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Human
US-09-801-191A-5

Query Match      31.9%; Score 46; DB 4; Length 576;
Best Local Similarity 36.7%; Pred. No. 42;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY      3 VVARLLHIGAIMFQRL-----DFIEQLSAPP 28
DB      187 VIARILHGGVMAQQQLLHVGDIIKEVNGQP 216

RESULT 8
US-09-328-352-7874
; Sequence 7874, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7874
; LENGTH: 1233
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7874

Query Match      31.6%; Score 45.5; DB 4; Length 1233;
Best Local Similarity 41.4%; Pred. No. 1.3e+02;
Matches 12; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

QY      1 HAVVARLLHIGAIMFOR-LDTIEQLSAPP 28
DB      486 HRSVADLVVEVDALFQIDFGEQVQYDP 514

RESULT 9
US-09-252-991A-24319
; Sequence 24319, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24319
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24319

Query Match      31.2%; Score 45; DB 4; Length 271;
Best Local Similarity 52.9%; Pred. No. 24;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      3 VVARLLHIGAIMFQRL 19
      ||||| : : : :
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Human
US-09-801-191A-5

Query Match      31.9%; Score 46; DB 4; Length 576;
Best Local Similarity 36.7%; Pred. No. 42;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY      3 VVARLLHIGAIMFQRL-----DFIEQLSAPP 28
DB      187 VIARILHGGVMAQQQLLHVGDIIKEVNGQP 216

RESULT 10
US-09-252-991A-31403
; Sequence 31403, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31403
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31403

Query Match      31.2%; Score 45; DB 4; Length 773;
Best Local Similarity 31.2%; Pred. No. 86;
Matches 5; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY      13 IMPQRLDFIEQLSAPP 28
DB      614 VVERQVEFVDQMGDPP 629

RESULT 11
US-08-567-508C-2
; Sequence 2, Application US/08567508C
; Patent No. 5914393
; GENERAL INFORMATION:
; APPLICANT: Coleman, Roger
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,508C
; FILING DATE: 05-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0049US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1132 amino acids
; TYPE: amino acid
```


Query Match 31.2%; Score 45; DB 3; Length 1139;
Best Local Similarity 25.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 2 AVVARLLHIGAIMFQRLDFEQLSAPPA 29
DB 1030 SVASDWSFGVVLIELFTYIEKSKSPPA 1057

RESULT 15
US-09-543-681A-6899
; Sequence 6899, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6899
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6899

Query Match 30.6%; Score 44; DB 4; Length 87;
Best Local Similarity 58.3%; Pred. No. 9;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 7 LLHIGAIMFQRL 18
DB 20 LVHLGMILFQRI 31

Search completed: November 10, 2004, 14:55:41
Job time : 10.6651 secs

This Page Blank (uspio)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	144	100.0	29	14	US-10-092-750-58		Sequence 58, Appl
2	51.5	35.8	176	15	US-10-424-599-153449		Sequence 153449,
3	49	34.0	112	17	US-10-425-115-204154		Sequence 204154,
4	49	34.0	392	16	US-10-437-963-163286		Sequence 163286,
5	48	33.3	318	15	US-10-425-114-64372		Sequence 64372, A
6	48	33.3	1008	14	US-10-369-493-5632		Sequence 5632, A
7	48	33.0	1008	14	US-10-369-493-5630		Sequence 5630, Ap
8	47.5	33.0	325	14	US-10-369-493-11896		Sequence 11896, A
9	47.5	33.0	623	16	US-10-437-963-127647		Sequence 127647,
10	47.5	33.0	645	15	US-10-437-963-127646		Sequence 127646,
11	47	32.6	795	14	US-10-456-761-9690		Sequence 9690, Ap
12	46.5	32.3	146	15	US-10-276-774-1599		Sequence 1599, Ap
13	46.5	32.3	184	14	US-10-028-072-76		Sequence 76, Appl

```

; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153449
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(176)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_109589C.1.pep
; US-10-424-599-153449

```

```

Query Match      35.8%; Score 51.5; DB 15; Length 176;
Best Local Similarity 35.7%; Pred. No. 7.9;
Matches 10; Conservative 9; Mismatches 8; Indels 1; Gaps 1;

QY 1 HAVVALLHIGAIMFQRLDFIEQLSAPP 28
| : : : : : : : : : : : : : : : :
Db 114 HPPIAQVLFYEVVLI-KLDFLERILAPP 140

```

```

RESULT 3
US-10-425-115-204154
; Sequence 204154, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 204154
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(112)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_117781C.1.pep
; US-10-425-115-204154

```

```

Query Match      34.0%; Score 49; DB 17; Length 112;
Best Local Similarity 39.3%; Pred. No. 12;
Matches 11; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

QY 1 HAVVALLHIGAIMFQRLDFIEQLSAPP 28
| : : : : : : : : : : : : : : : :
Db 39 HHAIVRLHLYSPFFPTLLFLVETKPP 66

```

```

RESULT 4
US-10-437-963-163286
; Sequence 163286, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

```

```

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163286
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62297C.1.pep
; US-10-437-963-163286

```

```

Query Match      34.0%; Score 49; DB 16; Length 392;
Best Local Similarity 44.4%; Pred. No. 47;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 12 AIMFQRLDFIEQLSAPPA 29
| : : : : : : : : : : : : : : : :
Db 347 AVVGDKVDFAEKLSPPS 364

```

```

RESULT 5
US-10-425-114-64372
; Sequence 64372, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64372
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-236-A4_FLI.pep
; US-10-425-114-64372

```

```

Query Match      33.3%; Score 48; DB 15; Length 318;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 HAVVALLHIGAIMFQRLDF 20
| : : : : : : : : : : : : : : : :
Db 268 HAVVDELEHLGALVRHGDG 287

```

```

RESULT 6
US-10-369-493-5629
; Sequence 5629, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

```

```
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5629
; LENGTH: 1008
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5629

Query Match      33.3%; Score 48; DB 14; Length 1008;
Best Local Similarity 27.0%; Pred. No. 1.9e+02;
Matches 10; Conservative 10; Mismatches 5; Indels 12; Gaps 1;

QY 3 VVARLLHIGA-----IMFQRLDFIEQLSAP 27
Db 303 LMASMLHIGAKDAKKEQEFELLDEKVDFTQALQMP 339

RESULT 7
US-10-369-493-5630
; Sequence 5630, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5630
; LENGTH: 1008
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5630

Query Match      33.3%; Score 48; DB 14; Length 1008;
Best Local Similarity 27.0%; Pred. No. 1.9e+02;
Matches 10; Conservative 10; Mismatches 5; Indels 12; Gaps 1;

QY 3 VVARLLHIGA-----IMFQRLDFIEQLSAP 27
Db 303 LMASMLHIGAKDAKKEQEFELLDEKVDFTQALQMP 339

RESULT 8
US-10-369-493-11896
; Sequence 11896, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
```

```
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11896
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-11896

Query Match      33.0%; Score 47.5; DB 14; Length 325;
Best Local Similarity 40.0%; Pred. No. 66;
Matches 10; Conservative 8; Mismatches 4; Indels 3; Gaps 1;

QY 5 ARLLHIGAIMFQRLDFIEQLSAPPA 29
Db 3 ARLVHIGSAV---VDYVYRIDALPA 24

RESULT 9
US-10-437-963-127647
; Sequence 127647, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127647
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_30078C.1.pep
US-10-437-963-127647

Query Match      33.0%; Score 47.5; DB 16; Length 623;
Best Local Similarity 46.7%; Pred. No. 1.3e+02;
Matches 14; Conservative 4; Mismatches 7; Indels 5; Gaps 2;

QY 1 HAVVARLLH--IGAIMFQRLDFIEQLSAPP 28
Db 189 HGRVAKLLRHITIGASV---VDFAYQLATPP 215

RESULT 10
US-10-437-963-127646
; Sequence 127646, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127646
```

```

; LENGTH: 645
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_30077C.1.pap
US-10-437-963-127646

Query Match      33.0%; Score 47.5; DB 16; Length 645;
Best Local Similarity 46.7%; Pred. No. 1.4e+02;
Matches 14; Conservative 4; Mismatches 7; Indels 5; Gaps 2;

QY 1 HAVARLLH--IGAINFORLDFIEQLSAPP 28
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 189 HGRVAKLRLHTIGASV---VDPAQLATPP 215

RESULT 11
US-10-156-761-9690
; Sequence 9690, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9690
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9690

Query Match      32.6%; Score 47; DB 14; Length 795;
Best Local Similarity 41.4%; Pred. No. 2.1e+02;
Matches 12; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 HAVARLLHIGAINFORLDFIEQLSAPP 29
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 225 HATVRLLDGVVDGLRIHDPGLADPDA 253

RESULT 12
US-10-276-774-1599
; Sequence 1599, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Yang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1599
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-276-774-1599

```

Query Match      32.3%; Score 46.5; DB 15; Length 146;
Best Local Similarity 48.3%; Pred. No. 39;
Matches 14; Conservative 3; Mismatches 9; Indels 3; Gaps 1;

QY 4 VARLLHIGAIMFOR---LDFIEQLSAPPA 29
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 88 VSALQHLKALDLSRNQFQDFPEQLTALPA 116

```

```

RESULT 13
US-10-028-072-76
; Sequence 76, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31

```

; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818

; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match 32.3%; Score 46.5; DB 14; Length 184;
Best Local Similarity 48.3%; Pred No. 50;
Matches 14; Conservative 3; Mismatches 9; Indels 3; Gaps 1;
QY 4 VARLLHIGAIMFQR---LDFIEQLSAPPA 29

Result No.	Query %		Length	DB	ID	Description
	Score	Match				
1	50.5	35.1	274	2	T03797	thymidylate kinase
2	50	34.7	740	2	D71602	hypothetical prote
3	50	34.7	1411	2	T18417	hypothetical prote
4	50	34.7	1417	2	T18418	hypothetical prote
5	48.5	33.7	742	2	S53663	hydrogenase-relate
6	48	33.3	192	2	A07058	precorrin-8W decar
7	48	33.3	342	2	S23438	hypothetical prote
8	48	33.3	610	2	T23452	hypothetical prote
9	48	33.3	651	2	AH3276	lysostaphin (EC 3 .
10	48	33.3	842	2	AB0674	probable hydrolase
11	48	33.3	1008	2	T19832	probable RNA helic
12	46	31.9	260	2	C82081	cell division prot
13	46	31.9	576	2	A57653	disks large homolo
14	45	31.2	275	2	C81700	3'-deoxy-manno-ocu
15	45	31.2	537	2	T41666	carl homolog - fis
16	45	31.2	1432	2	JW0091	Janus kinase (EC 2
17	45	31.2	1201	2	T08603	kinase-related pr
18	44.5	30.9	513	2	S21535	nitrate reductase
19	44	30.6	206	1	I40173	orotate phosphorib
20	44	30.6	215	2	E75058	hypothetical prote
21	44	30.6	221	1	A26753	glutathione transf
22	44	30.6	303	2	T23583	hypothetical prote
23	44	30.6	345	2	H95381	probable iron upta
24	44	30.6	469	2	G70699	probable rdaA prot
25	44	30.6	540	2	S74233	transcription fact
26	44	30.6	1515	2	S51824	myosin heavy chain
27	43.5	30.2	469	2	AD1926	hypothetical prote
28	43.5	30.2	505	2	S75722	UDP-N-acetylmuram
29	43.5	30.2	834	2	T19010	hypothetical prote

Matches 10; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* sero
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AB0674
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-842 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD01763.1; PID:gl6502611; GSPDB:GN00176
C;Genetics:
A;Gene: STY1504
C;Superfamily: maltotoligosyl trehalose synthase

Query Match 33.3%; Score 48; DB 2; Length 842;
Best Local Similarity 34.5%; Pred. No. 21;
Matches 10; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
QY 1 HAVVARLLHIGAIMFORLDFIEQLSAPPA 29
DB 233 HRLLELVRTGAVDGLRIDHIDGLADPKA 261

RESULT 11
T18932
probable RNA helicase - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18932; T18934
R;Percy, C.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19028
A;Accession: T18932
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1008 <WIL>
A;Cross-references: UNIPROT:O45244; EMBL:Z81457; PIDN:CAB03819.1; GSPDB:GN00020; CESP:C
A;Experimental source: clone C01G12
R;Lloyd, C.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19046
A;Accession: T18934
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1008 <WIL>
A;Cross-references: EMBL:Z81462; PIDN:CAB03845.1; GSPDB:GN00020; CESP:C04H5.6
C;Genetics:
A;Gene: CESP:C04H5.6
A;Map position: 2
A;Introns: 356/3; 441/3; 480/2; 635/3; 682/3; 822/2; 895/3; 923/1; 965/3; 983/3
C;Keywords: ATP; nucleotide binding; P-loop; pre-mRNA splicing
F;387-594/Region: nucleotide-binding motif A (P-loop)
F;481-486/Region: nucleotide-binding motif B
F;485-488/Region: DEAH motif

Query Match 33.3%; Score 48; DB 2; Length 1008;
Best Local Similarity 27.0%; Pred. No. 25;
Matches 10; Conservative 10; Mismatches 5; Indels 12; Gaps 1;
QY 3 VVARLLHIGA-----IMFORLDFIEQLSAP 27
DB 303 LMASMLHIGAKDAKKEQEFELLDDKVDPIQALQMP 339

RESULT 12
C82081
cell division protein FtsQ VC2399 [imported] - *Vibrio cholerae* (strain N16961 serogroup
C;Species: *Vibrio cholerae*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: C82081
R;Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000

Query Match 33.3%; Score 48; DB 2; Length 651;
Best Local Similarity 44.0%; Pred. No. 16;
Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;
QY 4 VARLLHIGAIMFORLDFIEQLSAPP 28
DB 44 VTSCMLGIALFAALDGREQLATPP 68

RESULT 10
AB0674
probable hydrolase STY1504 [imported] - *Salmonella enterica* subsp. *enterica* serovar Typh
C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A;Note: This species has also been called *Salmonella typhi*
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-May-2004
C;Accession: AB0674
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001

Query Match 33.3%; Score 48; DB 2; Length 610;
Best Local Similarity 47.8%; Pred. No. 15;
Matches 11; Conservative 4; Mismatches 6; Indels 2; Gaps 1;
QY 7 LHIGAIMFORLDFIEQLSAPPA 29
DB 222 ILYIGALCV--CDFVMSLSLPPA 242

RESULT 9
AH3276
lysostaphin (EC 3.5.1.1) [imported] - *Brucella melitensis* (strain 16M)
C;Species: *Brucella melitensis*
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AH3276
R;DeIvecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A;Reference number: AH3252; PMID:11756688
A;Accession: AH3276
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-651 <KUR>
A;Cross-references: UNIPROT:Q8FYU9; UNIPROT:Q8FYU0; GB:AE008917; PIDN:AA151379.1; PID:gl
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BME10197
A;Map position: 1
C;Keywords: hydrolase

Query Match 33.3%; Score 48; DB 2; Length 610;
Best Local Similarity 44.0%; Pred. No. 16;
Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;
QY 4 VARLLHIGAIMFORLDFIEQLSAPP 28
DB 44 VTSCMLGIALFAALDGREQLATPP 68

RESULT 10
AB0674
probable hydrolase STY1504 [imported] - *Salmonella enterica* subsp. *enterica* serovar Typh
C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A;Note: This species has also been called *Salmonella typhi*
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-May-2004
C;Accession: AB0674
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: C82081
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <HEI>
A;Cross-references: UNIPROT:Q9KPG9; GB:AE004310; GB:AE003852; NID:g9656963; PIDN:AAF9554
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2399
A;Map position: 1
C;Superfamily: cell division protein ftsQ

Query Match 31.9%; Score 46; DB 2; Length 260;
Best Local Similarity 42.9%; Pred. No. 12;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 3 VVARLHIGAIMFQRLDFIEQ 23
DB 74 VLARLDHIGTFMSQDINVLQE 94

RESULT 13

A57653
disks large homolog DLG2 - human
C;Species: Homo sapiens (man)
C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
C;Accession: A57653
R;Mazoyer, S.; Gayther, S.A.; Nagai, M.A.; Smith, S.A.; Dunning, A.; van Rensburg, E.J.;
Genomics 28, 25-31, 1995
A;Title: A gene (DLG2) located at 17q12-q21 encodes a new homologue of the Drosophila tu
A;Reference number: A57653; MUID:96070428; PMID:7590743
A;Accession: A57653
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-576 <MAZ>
A;Cross-references: UNIPROT:Q14168; GB:X82895; NID:g939884; PIDN:CAA58067.1; PID:g939885
C;Genetics:

A;Gene: GDB:DLG2; GDB:MPP2
A;Cross-references: GDB:6053914; GDB:567459; OMIM:600723
A;Map position: 17q12-17q21
F;170-239/Domain: GLGF domain homology <GLG>
F;263-311/Domain: SH3 homology <SH3>
F;375-564/Domain: guanylate kinase homology <GKI>

Query Match 31.9%; Score 46; DB 2; Length 576;
Best Local Similarity 36.7%; Pred. No. 29;
Matches 11; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

QY 3 VVARLHIGAIMFQRL---DFIEQLSAPP 28
DB 187 VIARLHGGWVAQQGLLVHGVGDIKEVNGQP 216

RESULT 14

C81700
3-deoxy-manno-octulosonate cytidylyltransferase TC0454 [imported] - Chlamydia muridarum
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
C;Accession: C81700
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: C81700
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-275 <TET>
A;Cross-references: GB:AE002314; GB:AE002160; NID:g7190495; PIDN:AAF39307.1; PID:g719049

A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0454

C;Superfamily: 3-deoxy-manno-octulosonate cytidylyltransferase

Query Match 31.2%; Score 45; DB 2; Length 275;
Best Local Similarity 40.9%; Pred. No. 19;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 8 LHIGAIMFQSLDFIEQLSAPPA 29
DB 200 LHIGVAFRRAPFLSEYVKIPPS 221

RESULT 15

T41666
Carl homolog - fission yeast (*Schizosaccharomyces pombe*)
C;Species: *Schizosaccharomyces pombe*
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T41666
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z22008
A;Accession: T41666
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-537 <WOO>
A;Cross-references: UNIPROT:O59833; EMBL:AL023590; PIDN:CAA19073.1; GSPDB:GN00068; SPDB
A;Experimental source: strain 972h-; cosmid c965
C;Genetics:
A;Gene: SPDB:SPCC965.13
A;Map position: 3
C;Superfamily: benomyl/methotrexate resistance protein

Query Match 31.2%; Score 45; DB 2; Length 537;
Best Local Similarity 47.1%; Pred. No. 39;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 AVVARLHIGAIMFQRL 18
DB 466 AIAAIVLHGLGIMFDNM 482

Search completed: November 10, 2004, 14:52:26
Job time : 8.11321 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 38.6302 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-58

Perfect score: 144
Sequence: 1 HAVVARLLHIGAIMFORLDFIEQLSAPPA 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	36.1	805	2 Q9ADI2	Q9adi2 streptomyc
2	51	35.4	77	2 Q8A2U7	Q8a2u7 bacteroides
3	50.5	35.1	274	2 Q60970	O60970 leishmania
4	50	34.7	1416	2 Q73309	Q73309 plasmodium
5	50	34.7	1417	2 Q73310	Q73310 plasmodium
6	50	34.7	1440	2 Q96279	Q96279 plasmodium
7	49.5	34.4	259	2 Q6XL72	Q6xl72 vitis vinif
8	49.5	34.4	259	2 AAP49697	Aap49697 vitis vin
9	49	34.0	203	2 Q9FW86	Q9fw86 oryza sativ
10	49	34.0	402	2 Q8PCD6	Q8pcd6 xanthomonas
11	49	34.0	481	1 TRNE_RALSO	Q8y3hs ralstonia s
12	49	34.0	804	2 Q7XQG9	Q7xqg9 oryza sativ
13	49	34.0	1296	2 Q8WRV3	Q8wrv3 plasmodium
14	48.5	33.7	155	2 Q88EQ7	Q88eq7 pseudomonas
15	48.5	33.7	198	2 Q6N7G4	Q6n7g4 rhodopsedu
16	48.5	33.7	198	2 Q6N7G4	Q6n7g4 rhodopsedu
17	48.5	33.7	342	1 HUPV_AZOC	Cae27734 rhodopseu
18	48.5	33.7	1022	2 Q8NYG8	Q43959 azotobacter
19	48.5	33.7	1022	2 A8H6602	Q6nyg8 brachydanio
20	48	33.3	192	1 CBIT_SALTI	A8h6602 brachydan
21	48	33.3	195	2 Q9C0Q3	Q825m9 salmonella
22	48	33.3	342	1 HUPK_AZOV	Q9c0q3 nectria hae
23	48	33.3	348	1 H0XV_AZOV	P31878 azotobacter
24	48	33.3	427	2 Q8EQS4	P40597 azotobacter
25	48	33.3	433	2 Q45725	Q8eqs4 oceanobacil
26	48	33.3	651	2 Q8FYLO	O45725 caenorhabdi
27	48	33.3	651	2 Q8YJ89	Q8fylo brucella su
28	48	33.3	842	2 Q8Z723	Q8yj89 brucella me
29	48	33.3	842	2 Q8ZPF1	Q8z723 salmonella
30	48	33.3	873	2 Q6N513	Q8zpf1 salmonella
31	48	33.3	873	2 CAE28611	Q6n513 rhodopsedu
					Cae28611 rhodopseu

Q9VYJ0 drosophila
O45244 caenorhabdi
Q7M9V7 woliniella s
Q98JW9 rhizobium l
Q880S1 pseudomonas
Q850Z4 oryza sativ
Q7VIX3 helicobacte
Q6FX09 candida gla
Q8WPU5 trypanosoma
Q87SG2 vibrio para
Q8KAX6 chlorobium
Q6NJE0 corynebacte
Cae48969 corynebac
Q82L60 streptomyce

ALIGNMENTS

RESULT 1

Q9ADI2 PRELIMINARY; PRT; 805 AA.
ID Q9ADI2
AC Q9ADI2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative alpha amylase.
GN OrderedLocusNames=SC06081; ORFNames=SC6AC1A6.05c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
Huang C.-H., Kieser T., Larke L., Murphree L.D., Oliver K., O'Neill S.,
Rabbinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)".
RL Nature 417:141-147(2002).
DR EMBL; AJ939126; CAC33926.1; -.
DR HSP; Q53688; 11V8.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR000408; Reg_ch_r_condens.
DR InterPro; IPR000408; Reg_ch_r_condens.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PROSITE; PS00626; RCCL_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 805 AA; 87203 MW; 6FAF921CA46AE13B CRC64;

Query Match 36.1%; Score 52; DB 2; Length 805;
Best Local Similarity 44.8%; Pred. No. 29;
Matches 13; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Qy 1 HAVVARLLHIGAIMFORLDFIEQLSAPPA 29

Db 229 HGTVRLHGBVIGLRLVDPGLADPDA 257

RESULT 2

Q8A2U7 PRELIMINARY; PRT; 77 AA.
ID Q8A2U7
AC Q8A2U7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BT3208;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AB016939; AAC78314.1; -.
DR InterPro; IPR001452; SH3.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 77 AA; 9146 MW; 6F6D49C357A0F215 CRC64;

Query Match 35.4%; Score 51; DB 2; Length 77;
Best Local Similarity 52.6%; Pred. No. 3.7;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 VVARLLHIGAIMFQRLDFI 21
DB 12 VVAVLLSLGALFFQLNFV 30

RESULT 3
O60970 PRELIMINARY; PRT; 274 AA.
AC O60970;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TKRPI.
GN Name=TKRPI;
OS Leishmania major.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=99178987; PubMed=10077609;
RA Myler P.J., Audleman L., deVos T., Hixson G., Kiser P., Lemley C.,
RA Magness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
RA Bastien P., Fu G., Ivens A., Stuart K.;
RT "Leishmania major Friedlin chromosome 1 has an unusual distribution of
RT protein-coding genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
DR EMBL; AB001274; AAC24621.1; -.
DR PIR; A81456; T02797.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004798; F:thymidylate kinase activity; IEA.
DR GO; GO:0004233; P:tdp biosynthesis; IEA.
DR GO; GO:0008235; P:tdp biosynthesis; IEA.
DR InterPro; IPR000062; Thymidylate_kin.
DR Pfam; PF02223; Thymidylate_kin; 1.
SQ SEQUENCE 274 AA; 29782 MW; B27AE442261FF7D3 CRC64;

Query Match 35.1%; Score 50.5; DB 2; Length 274;
Best Local Similarity 44.1%; Pred. No. 17;
Matches 15; Conservative 5; Mismatches 7; Indels 7; Gaps 2;

QY 2 AVVARLLHIGAIMFQRLD-----FIQLSAPP 28
DB 5 AALARTLHGAGNVYRSLDCLKVLAVQSSPPP 38

RESULT 4
O77309 PRELIMINARY; PRT; 1416 AA.

```

```

ID O77309 PRELIMINARY; PRT; 1416 AA.
AC O77309; O77307;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytoadherence linked asexual protein, CLAG.
GN Name=MAJ3P1.2; Synonyms=PFC0110w;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
[1]
RN SEQUENCE FROM N.A.
RC MEDLINE=99376085; PubMed=10448855;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
RA Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jaseal B., Kyes S., McLean J., Moule S.,
RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
RT falciparum.";
RL Nature 400:532-538(1999).
RN [2]
RN SEQUENCE FROM N.A.
RC MEDLINE=2255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C., Corton C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Duggan J.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; Z97348; CAB10571.3; -.
DR PIR; T18417; T18417.
DR InterPro; IPR005553; CLAG.
DR Pfam; PF03805; CLAG; 1.
SQ SEQUENCE 1416 AA; 167488 MW; 1E15EF2D1CD91BB5 CRC64;

Query Match 34.7%; Score 50; DB 2; Length 1416;
Best Local Similarity 45.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 7 LLHIGAIMFQRLDFIQLSAPP 28
DB 691 LVHVS--FLQDFFHQLNEPP 710

RESULT 5
O77310 PRELIMINARY; PRT; 1417 AA.
AC O77310; O77311;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytoadherence linked asexual protein, CLAG.
GN Name=MAJ3P1.5; Synonyms=PFC0120w;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
[1]
RN SEQUENCE FROM N.A.
RC MEDLINE=99376085; PubMed=10448855;

```

```

Bowman S., Lawson D., Baeham D., Brown D., Chillingworth T.,
Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moule S.,
Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
Taylor K., Tivey A., Whitehead S., Woodhead S., Woodward J.,
Whithead S., Woodward J.R., Newbold C., Barrall B.G.;
"the complete nucleotide sequence of chromosome 3 of Plasmodium
falciiparum.";
RL Nature 400:532-538 (1999).
[2]
SEQUENCE FROM N.A.
MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabbintowitsch E.,
Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrall B.G.;
RT "Sequence of Plasmodium falciiparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531 (2002).
DR EMBL; Z97348; CAB10572.2; -.
DR PIR; T18418.
DR InterPro; IPR005553; CLAG.
DR Pfam; PF03805; CLAG; 1.
SQ SEQUENCE 1417 AA; 167241 MW; AE7AA0E54B5B1130 CRC64;

Query Match 34.7%; Score 50; DB 2; Length 1417;
Best Local Similarity 45.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 7 LLHGCAIMFQELDFIEQLSAPP 28
Db 689 LVHVNYS--FLQDFFHQLNEPP 708

RESULT 6
ID O96279 PRELIMINARY; PRT; 1440 AA.
AC O96279; O96280;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytoadherence linked asexual protein 2.
GN Name=PF0935w;
OS Plasmodium falciiparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
[1]
SEQUENCE FROM N.A.
MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shallom S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
falciiparum.";
RL Science 282:1126-1132 (1998).
[2]
SEQUENCE FROM N.A.
MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,

```

```

RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyse S.,
RA Chan M.-S., Nene V., Shalloo S.J., Sun B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Murgall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrall B.;
RA "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RT Nature 419:498-511(2002).
RL EMBL; AB001428; AAC71977.2; -.
DR PIR; C71602; C71602.
DR PIR; D71602; D71602.
DR InterPro; IPR005553; CLAG.
DR InterPro; IPR002160; Prot_inh_Kunz-1g.
DR Pfam; PF03805; CLAG; 1
DR PROSITE; PS00283; SOYBEAN_KUNITZ; UNKOWN_1.
SQ SEQUENCE 1440 AA; 171159 MW; 876825442EBC66F9 CRC64;

Query Match 34.7%; Score 50; DB 2; Length 1440;
Best Local Similarity 45.5%; Pred.No. 1.le+02;
Matches 10; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 7 LHIGAIMFORLDIEQLSAPP 28
Db 716 LVHVS--FLQLDFHQLNEPP 735
|.: : |::| ::||
:-: -:-: -:-: -:-:

RESULT 7
ID Q6XL72 PRELIMINARY; PRT; 259 AA.
AC Q6XL72;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome P-450-like protein (Fragment).
GN Names=CYP2;
OS Vitis vinifera (Grape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Vitaceae; Vitis.
OX NCBI_TaxID=29760;
ON [1]
RN SEQUENCE FROM N.A.
RP Bezier A., Lambert B., Baillieul F.;
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AY226829; AAP49697.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002401; EP450I.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00463; EP450I.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR FT Name; Monooxygenase; Oxidoreductase.
KW NON TER
FT 1
SQ SEQUENCE 259 AA; 28776 MW; 261C3D9532355817 CRC64;

Query Match 34.4%; Score 49.5; DB 2; Length 259;
Best Local Similarity 41.2%; Pred.No. 23;
Matches 14; Conservative 5; Mismatches 10; Indels 5; Gaps 2

QY 1 HAVVARLLH---IGAIMFORLDIEQ--LSAPPA 29
Db 205 HLTLRIHLHFELGAVADSPVDMTESGLTAPKA 238
|.: : |::| ::||
:-: -:-: -:-: -:-:

RESULT 8
AAP49697 PRELIMINARY; PRT; 259 AA.
ID AAP49697
AC AAP49697;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
```

DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
 DE Cytochrome P-450-like protein (Fragment).
 GN CYP82.
 OS Vitis vinifera (Grape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Vitaceae; Vitis.
 OX NCBI_TaxID=29760;
 RN [1]
 RZ SEQUENCE FROM N.A.
 RC STRAIN=cv. Chardonnay;
 RA Bezier A., Lambert B., Baillieux F.;
 RT "Molecular cloning of partial cytochrome P450-like protein mRNA from
 RT grapevine."
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY226829; AAP49697.1; -
 FT NON TER 1 1
 SQ SEQUENCE 259 AA; 28776 MW; 261C3D9532355817 CRC64;

Query Match 34.4%; Score 49.5; DB 2; Length 259;
 Best Local Similarity 41.2%; Pred. No. 23;
 Matches 14; Conservative 5; Mismatches 10; Indels 5; Gaps 2;

QY 1 HAVVARELLH---IGATMFORLDFIEQ--LSAPPA 29
 Db 205 HLTALRLHGFELGAVDSFVDMTESPGLTAPKA 238

RESULT 9

Q9FW86 PRELIMINARY; PRT; 203 AA.
 AC Q9FW86; Q7XC2;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DE 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Putative cytochrome P450.
 GN ORFNames=OSJNBa0026L2.11;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RZ SEQUENCE FROM N.A.
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
 RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RZ SEQUENCE FROM N.A.
 RA Buell R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RZ SEQUENCE FROM N.A.
 RA The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10."
 RL Science 300:1566-1569 (2003).
 RN [4]
 RZ SEQUENCE FROM N.A.
 RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 DR EMBL: AC068924; AG13497.1; -
 DR EMBL: AE017112; AAP54605.1; -
 DR HSSP: P11712; LOG2.
 DR Gramene; Q7XC2; -
 DR Gramene; Q9FW86; -
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR002401; EP450I.

DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00463; EP450I.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 203 AA; 22990 MW; 859FFD62EB372645 CRC64;

Query Match 34.0%; Score 49; DB 2; Length 203;
 Best Local Similarity 44.4%; Pred. No. 21;
 Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 12 AIMFQRLDFIEQLSAPPA 29
 Db 158 AVGDKVDFAEKLSPPS 175

RESULT 10

Q8PCD6 PRELIMINARY; PRT; 402 AA.
 ID Q8PCD6
 AC Q8PCD6;
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE 2-octaprenyl-6-methoxyphenol hydroxylase.
 GN Name=ubiH; OrderedLocustNames=XCC0798;
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RZ SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Cimarotte G., Cannavan F., Cardozo J., Chamberg F., Giapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Fatsuyama E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.P.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities."
 RL Nature 417:459-463 (2002).
 DR EMBL; AE012179; AAM40113.1; -
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR000733; Flav monooxygenase.
 DR InterPro; IPR000205; NAD_ES.
 DR InterPro; IPR003042; Kmg_mnoxygenase.
 DR InterPro; IPR011295; UbiH.
 DR InterPro; IPR010971; Ubi_Hases.
 DR Pfam; PF01360; Monooxygenase; 1.
 DR PRINTS; PR00420; NONGMNOXGNASE.
 DR TIGRfams; TIGR01988; Ubi-Hases; 1.
 DR TIGRfams; TIGR01984; UbiH; 1.
 KW complete proteome.
 SQ SEQUENCE 402 AA; 42578 MW; AC80895FF3CCBF31 CRC64;

Query Match 34.0%; Score 49; DB 2; Length 402;
 Best Local Similarity 37.1%; Pred. No. 43;
 Matches 13; Conservative 5; Mismatches 7; Indels 10; Gaps 2;

QY 1 HAVVARELLHIGATMFORLDFIEQ-----SAPP 28
 Db 1 HAVVARELLHIGATMFORLDFIEQ-----SAPP 28

```

Db      169 HSAVRELLHGT---EQHDFLQTLFVARVRASEPP 200

RESULT 11
TRME RALSO STANDARD; PRT; 481 AA.
AC QBYJH5;
AD 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DT NAME=trmE; Synonyms=thdF; OrderedLocusNames=RSC0005; ORFNames=RS01827;
DN Ralstonia solanacearum (Pseudomonas solanacearum).
OS Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
CC NCBI_TaxID=305;
XX [1]
RN SEQUENCE FROM N.A.
RR STRAIN=GMI1000;
RC MEDLINE=21618179; PubMed=11923852; DOI=10.1038/415497a;
RX Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RT Nature 415:497-502(2002).
RL -I- FUNCTION: Exhibits a very high intrinsic GTPase hydrolysis rate.
CC Involved in the biosynthesis of the hypermodified nucleoside 5-methylaminomethyl-2-thiouridine, which is found in the wobble position of some tRNAs (By similarity).
CC -I- SIMILARITY: Belongs to the era/trmE family of GTP-binding proteins. trmE subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
-----
CC ENBL; AL646057; CAD13533.1; -.
DR HAMAP; MF 00379; -, 1.
DR InterPro; IPR005289; GTP-binding.
DR InterPro; IPR006073; Gtp1_OBG.
DR InterPro; IPR001806; Ras_Transfrrng.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR004520; thdF.
DR PRINTS; PR00326; GTF1OBG.
DR PRINTS; PR00449; RASTRNSFRMG.
DR TIGRFAMs; TIGR00650; MG442; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR TIGRFAMs; TIGR00450; thdF; 1.
KW Complete proteome; GTP-binding; tRNA processing.
FT NP_BIND 245 252 GTP (Potential).
FT NP_BIND 292 296 GTP (Potential).
FT NP_BIND 361 364 GTP (Potential).
SQ SEQUENCE 481 AA; 50908 MW; 0D54949EE2B7E54D CRC64;

Query Match 34.0%; Score 49; DB 1; Length 481;
Best Local Similarity 39.1%; Pred.No.52;
Matches 9; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HAVVARLIHGAIINFORLDFFIEQ 23
       |||::||: :: |||: |
Db      179 HALVERVIHLRMVLVEATLDFFPE 101

RESULT 12
ID Q7XQG9 PRELIMINARY; PRT; 804 AA.
```

Db 654 VQYVIHVS--FMQLDFFHALNEPP 677

RESULT 14

Q88EQ7 PRELIMINARY; PRT; 155 AA.

AC Q88EQ7; (Created)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Hypothetical protein.

GN OrderedLocNames=PP4396;

OS Pseudomonas putida (strain KT2440).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=160488;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22423060; PubMed=12534453;

RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,

RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,

RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,

RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,

RA Hance I., Chris Lee P., Holtzaple E.K., Scanlan D., Tran K.,

RA Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,

RA Wedler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,

RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,

RA Fraser C.M.;

RT "Complete genome sequence and comparative analysis of the

RT metabolically versatile Pseudomonas putida KT2440.";

RL Environ. Microbiol. 4:799-808(2002).

DR EMBL; AS016720; AN69974.1; -.

DR TTGR; PF4396; -.

DR GO; GO:0019861; C:flagellum; IEA.

DR GO; GO:0009296; P:flagellum biogenesis; IEA.

DR InterPro; IPR007809; FlgN.

DR Pfam; PF05130; FlgN; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 155 AA; 17170 MW; 9070C76F8C705C47 CRC64;

Query Match 33.7%; Score 48.5; DB 2; Length 155;

Best Local Similarity 52.2%; Pred. No. 19;

Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 2 AVVARLLHIGAIMFQRLDFIEQL 24

DB 78 AVAAQSPH-GDVMLOKLDMLSQL 99

RESULT 15

Q6N7G4 PRELIMINARY; PRT; 198 AA.

AC Q6N7G4;

DT 05-JUL-2004 (TReMBLrel. 27, Created)

DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)

DE Hypothetical protein.

GN OrderedLocNames=RPA2293;

OS Rhodopsseudomonas palustris.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae; Rhodopsseudomonas.

OX NCBI_TaxID=1076;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CGA009 / ATCC BAA-98;

RX PubMed=14704707; DOI=10.1038/nbt923;

RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,

RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,

RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,

RA Harrison F.H., Gibson J., Harwood C.S.;

RT "Complete genome sequence of the metabolically versatile

RT photosynthetic bacterium Rhodopsseudomonas palustris.";

RL Nat. Biotechnol. 22:55-61(2004).

DR EMBL; BX572600; CAB27734.1; -.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 198 AA; 22169 MW; DCDB2421AE73DB9B CRC64;

Query Match 33.7%; Score 48.5; DB 2; Length 198;

Best Local Similarity 48.3%; Pred. No. 25;

Matches 14; Conservative 2; Mismatches 10; Indels 3; Gaps 1;

QY 1 HAVVARLLHIGAIMFQRLDFIEQLSAPPA 29

DB 44 HRAVAELLISRTIFAERLD---ALSAHPS 69

Search completed: November 10, 2004, 14:50:26

Job time : 40.6302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 11.3472 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-59

Perfect score: 178
Sequence: 1 GCGTLGSGMEAWLATVLKALPWHPTYQLEP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCFUS COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51.5	28.9	644	US-09-107-532A-7277	Sequence 7277, Ap
2	51	28.7	441	US-09-543-681A-7207	Sequence 7207, Ap
3	50.5	28.4	57	US-09-513-999C-6363	Sequence 6363, Ap
4	50.5	28.4	60	US-09-270-767-34498	Sequence 34498, A
5	50.5	28.4	60	US-09-270-767-49715	Sequence 49715, A
6	50	28.1	308	US-09-489-039A-14278	Sequence 14278, A
7	49.5	27.8	502	US-09-724-737-90	Sequence 90, Appl
8	49	27.5	413	US-08-282-197C-49	Sequence 49, Appl
9	49	27.5	446	US-09-543-681A-5864	Sequence 5864, Ap
10	49	27.5	557	US-09-413-814-34	Sequence 34, Appl
11	48.5	27.2	345	US-09-489-039A-7417	Sequence 7417, Ap
12	48.5	27.2	1239	US-09-252-991A-30198	Sequence 30198, A
13	48	27.0	208	US-09-248-796A-24595	Sequence 24595, A
14	48	27.0	326	US-09-489-039A-13604	Sequence 13604, A
15	47.5	26.7	80	US-09-252-991A-22806	Sequence 22806, A
16	47	26.4	29	US-08-293-778-6	Sequence 6, Appl
17	47	26.4	138	US-07-857-224B-112	Sequence 112, App
18	47	26.4	151	US-08-614-935-28	Sequence 28, Appl
19	47	26.4	151	US-09-130-287-28	Sequence 28, Appl
20	47	26.4	211	US-09-205-258-1159	Sequence 1159, Ap
21	47	26.4	217	US-09-205-258-1162	Sequence 1162, Ap
22	47	26.4	243	US-09-541-759-7	Sequence 7, Appl
23	47	26.4	259	US-08-944-483-52	Sequence 52, Appl
24	47	26.4	259	US-10-165-442-2	Sequence 2, Appl
25	47	26.4	259	US-10-165-442-4	Sequence 4, Appl
26	47	26.4	295	US-08-338-368-2	Sequence 2, Appl
27	47	26.4	295	US-10-165-442-1	Sequence 1, Appl

28 47 26.4 295 4 US-10-165-442-3 Sequence 3, Appl
29 47 26.4 376 2 US-08-558-269-10 Sequence 10, Appl
30 47 26.4 376 3 US-09-410-882-10 Sequence 10, Appl
31 47 26.4 416 4 US-09-489-039A-13142 Sequence 13142, A
32 47 26.4 579 1 US-08-295-411-4 Sequence 4, Appl
33 47 26.4 579 2 US-08-955-471-4 Sequence 4, Appl
34 47 26.4 579 3 US-09-117-708-14 Sequence 14, Appl
35 47 26.4 579 5 PCT-US92-10242-4 Sequence 4, Appl
36 47 26.4 615 1 US-07-998-972A-3 Sequence 3, Appl
37 47 26.4 615 1 US-08-463-953-3 Sequence 3, Appl
38 47 26.4 615 1 US-08-462-261-3 Sequence 3, Appl
39 47 26.4 615 5 PCT-US92-11357-3 Sequence 3, Appl
40 47 26.4 622 3 US-08-952-967-8 Sequence 8, Appl
41 47 26.4 622 4 US-09-054-272-42 Sequence 42, Appl
42 46.5 26.1 93 4 US-09-461-325-176 Sequence 176, App
43 46.5 26.1 93 4 US-10-012-542-176 Sequence 176, App
44 46.5 26.1 93 4 US-10-115-123-176 Sequence 176, App
45 46.5 26.1 129 4 US-09-461-325-379 Sequence 379, App

ALIGNMENTS

RESULT 1
US-09-107-532A-7277
; Sequence 7277, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS: 7310
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 7277:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...644
SEQUENCE DESCRIPTION: SEQ ID NO: 7277:
US-09-107-532A-7277

```
Query Match      28.4%; Score 50.5; DB 4; Length 57;
Best Local Similarity 36.7%; Pred. No. 5.3;
Matches 11; Conservative 4; Mismatches 8; Indels 7; Gaps 1;
```

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCES: 2709, 2004001
 CURRENT APPLICATION NUMBER: US/09/489, 039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0410000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 413 amino acids
; TYPE: amino acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-282-197C-49

Query Match 27.5%; Score 49; DB 2; Length 413;
Best Local Similarity 80.0%; Pred. No. 88;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0

QY 20 ALPWHPTYQL 29
Db 146 ALVWHPSYQL 155

RESULT 9
US-09-543-681A-5864
; Sequence 5864, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEIN
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5864
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5864

Query Match 27.5%; Score 49; DB 4; Length 446;
Best Local Similarity 47.8%; Pred. No. 96;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0

QY 3 GTLNGSGMEAWLATVILKALPWP 25
Db 321 GILKSGLSDSLAVILSLNDMHP 343

RESULT 10
US-09-413-814-34
; Sequence 34, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloeker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Kueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
; TITLE OF INVENTION: heteropolypeptide compounds
; FILE REFERENCE: PCT/US 99/23535

```

; CURRENT APPLICATION NUMBER: US/09/413,814
 ; CURRENT FILING DATE: 1999-10-07
 ; EARLIER APPLICATION NUMBER: DE 198 46 493.2
 ; EARLIER FILING DATE: 1998-10-09
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 34
 ; LENGTH: 557
 ; TYPE: PRT
 ; ORGANISM: Sorangium cellulosum
 US-09-413-814-34

Query Match 27.5%; Score 49; DB 3; Length 557;
 Best Local Similarity 47.6%; Pred. No. 1.2e+02;
 Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 5 LMGSGMEAWLATVLKALPWP 25
 Db 491 LHGLDPEAYLADVIRAMPYWP 511

RESULT 11
 US-09-489-039A-7417
 ; Sequence 7417, Application US/09489039A
 ; Patent No. 6610836

; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 7417
 ; LENGTH: 345
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-7417

Query Match 27.2%; Score 48.5; DB 4; Length 345;
 Best Local Similarity 42.3%; Pred. No. 84;
 Matches 11; Conservative 2; Mismatches 10; Indels 3; Gaps 1;

Qy 1 GQGLMGSGMEA---WLATVLKALPW 23
 Db 29 GRAPLMGEGLLAALWMLGGLFTLTW 54

RESULT 12
 US-09-252-991A-30198
 ; Sequence 30198, Application US/09252991A
 ; Patent No. 6551795

; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 30198
 ; LENGTH: 1239
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-30198

Query Match 27.2%; Score 48.5; DB 4; Length 1239;

Best Local Similarity 50.0%; Pred. No. 3.7e+02;
 Matches 9; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

Qy 6 WGSMEAWLATVLKALPW 23
 Db 15 W-SQIAWLVMVWPWP 31

RESULT 13
 US-09-248-796A-24595
 ; Sequence 24595, Application US/09248796A
 ; Patent No. 6747137

; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 24595
 ; LENGTH: 208
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 US-09-248-796A-24595

Query Match 27.0%; Score 48; DB 4; Length 208;
 Best Local Similarity 44.4%; Pred. No. 54;
 Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 3 GTLWGSMEAWLATVLKA 20
 Db 1 GFFWGGGFPWAGTALVA 18

RESULT 14
 US-09-489-039A-13604
 ; Sequence 13604, Application US/09489039A
 ; Patent No. 6610836

; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 13604
 ; LENGTH: 326
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-13604

Query Match 27.0%; Score 48; DB 4; Length 326;
 Best Local Similarity 40.7%; Pred. No. 92;
 Matches 11; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

Qy 4 TLMGSGMEAWLATVLKALPWHPTTYQL 30
 Db 47 TLACGSSTAWRKDAVLAVLPQTLOOB 73

RESULT 15
 US-09-252-991A-22806
 ; Sequence 22806, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

Fri Nov 12 14:55:45 2004

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22806
; LENGTH: 80
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22806

```

```

Query Match      26.7%; Score 47.5; DB 4; Length 80;
Best Local Similarity 57.9%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 5 LWMGMEAE---WLATVLKA 20
   |:|:|:|:|:|:|
DB 15 LFGGMEAEARFWLSLVLSA 33

```

Search completed: November 10, 2004, 14:55:42
Job time : 12.3972 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 35.8547 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-59

Perfect score: 178

Sequence: 1 GQGTWGSMEAWLATVLPKALPHPTVQLEP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	178	100.0	31	14	US-10-092-750-59
2	62.5	35.1	241	16	Sequence 59, Appli
3	60.5	34.0	94	15	Sequence 135353,
4	60.5	34.0	526	16	Sequence 3360, Ap
5	60	33.7	131	17	Sequence 164238
6	58.5	31.2	569	15	Sequence 271630,
7	55.5	31.2	584	15	Sequence 279403,
8	55	30.9	853	16	Sequence 124083,
9	54.5	30.6	48	9	Sequence 102, App
10	54.5	30.6	242	17	Sequence 235453,
11	54	30.3	74	17	Sequence 135353,
12	54	30.3	575	15	Sequence 228082,
13	54	30.3	575	15	Sequence 62303, A
					Sequence 64407, A

14	53.5	30.1	542	15	US-10-425-114-69836	Sequence 69836, A
15	53.5	30.1	607	17	US-10-425-115-198410	Sequence 198410,
16	53.5	30.1	618	15	US-10-425-114-64802	Sequence 64802, A
17	53	29.8	11300	16	US-10-250-304A-2	Sequence 2, Appli
18	52.5	29.5	55	9	US-09-864-761-40496	Sequence 40496, A
19	52.5	29.5	63	9	US-09-864-761-41691	Sequence 41691, A
20	52	29.2	131	17	US-10-425-115-362054	Sequence 362054,
21	52	29.2	218	15	US-10-425-114-53728	Sequence 52728, A
22	52	29.2	335	17	US-10-425-115-208290	Sequence 208290,
23	52	29.2	350	15	US-10-425-114-58750	Sequence 58750, A
24	51.5	28.9	50	9	US-09-864-761-40675	Sequence 40675, A
25	51.5	28.9	224	14	US-10-104-047-2601	Sequence 2601, Ap
26	51.5	28.9	599	16	US-10-437-963-154605	Sequence 154605,
27	51.5	28.9	763	13	US-10-081-408-2	Sequence 2, Appli
28	51.5	28.9	998	13	US-10-081-408-20	Sequence 20, Appli
29	51	28.7	130	15	US-10-424-599-266380	Sequence 266380,
30	51	28.7	213	17	US-10-425-115-292104	Sequence 292104,
31	51	28.7	214	15	US-10-282-122A-68380	Sequence 68380, A
32	51	28.7	312	16	US-10-437-963-132094	Sequence 132094,
33	51	28.7	410	15	US-10-425-114-38092	Sequence 38092, A
34	51	28.7	455	17	US-10-425-115-246019	Sequence 246019,
35	51	28.7	801	15	US-10-389-566-2305	Sequence 2305, Ap
36	51	28.7	863	17	US-10-425-115-307525	Sequence 307525,
37	50.5	28.4	150	15	US-10-424-599-257786	Sequence 257786,
38	50.5	28.4	567	16	US-10-437-963-183584	Sequence 183584,
39	50.5	28.4	981	14	US-10-174-677-31	Sequence 31, Appli
40	50.5	28.4	981	15	US-10-380-705-2	Sequence 88, Appli
41	50.5	28.4	981	15	US-10-210-172-88	Sequence 218149,
42	50	28.1	74	15	US-10-424-599-218149	Sequence 218149,
43	50	28.1	88	16	US-10-437-963-145388	Sequence 145388,
44	50	28.1	111	15	US-10-424-599-145125	Sequence 145125,
45	50	28.1	229	14	US-10-017-161-2148	Sequence 2148, Ap

ALIGNMENTS

RESULT 1
US-10-092-750-59
; Sequence 59, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-59

Query Match 100.0%; Score 178; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 6.3e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GQGTWGSMEAWLATVLPKALPHPTVQLEP 31
DB 1 GQGTWGSMEAWLATVLPKALPHPTVQLEP 31
RESULT 2
US-10-437-963-135353
; Sequence 135353, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

```

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135353
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(241)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37038C.1.pbp
; US-10-437-963-135353

Query Match      35.1%; Score 62.5; DB 16; Length 241;
Best Local Similarity 47.1%; Pred. No. 5.1;
Matches 16; Conservative 4; Mismatches 11; Indels 3; Gaps 2;

QY      1 GGGTGWGSGMEAWLATVVKALP--WHPTYQLEP 31
DB      32 GGGVWRCGSGGGLAGVRRSLPYLWPPPTQLPP 65

RESULT 3
US-10-264-049-3360
; Sequence 3360, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3360
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (76)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (93)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; US-10-264-049-3360

```

```

Query Match      34.0%; Score 60.5; DB 15; Length 94;
Best Local Similarity 45.8%; Pred. No. 3.6;
Matches 11; Conservative 1; Mismatches 7; Indels 5; Gaps 1;

QY      1 GGGTGWGSGMEAWLATVVKALPWH 24
DB      38 GRVTFWGCWEAWC-----LLRWH 56

RESULT 4
US-10-437-963-164238
; Sequence 164238, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 164238
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63157C.1.pbp
; US-10-437-963-164238

```

```

Query Match      34.0%; Score 60.5; DB 16; Length 526;
Best Local Similarity 38.1%; Pred. No. 20;
Matches 16; Conservative 5; Mismatches 6; Indels 15; Gaps 3;

QY      1 GGGTGWGSGM-----EA--WLATVVKALPWH-----PTY 27
DB      252 GEGSQSGRLLLPPEANMTALVIEASPMWRRTDVAVPYPT 293

RESULT 5
US-10-425-115-271630
; Sequence 271630, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 271630
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_17931C.1.pbp
; US-10-425-115-271630

Query Match      33.7%; Score 60; DB 17; Length 131;
Best Local Similarity 38.7%; Pred. No. 5.8;
Matches 12; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

QY      1 GGGTGWGSGMEAWLATVVKALPWHPTYQLEP 31

```


Db 22 GGRRWAVG--SFVVTIMGATPWLPPRQSAP 50

RESULT 6

US-10-424-599-279403
; Sequence 279403, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 279403

; LENGTH: 569

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_94322C.1.pep
US-10-424-599-279403

Query Match 31.2%; Score 55.5; DB 15; Length 569;

Best Local Similarity 52.0%; Pred. No. 98;

Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 2 GQTLWGSQMEAWL-ATVLKALPWHPP 25

Db 121 EGTLMWHAHSSWLKATVYCALIHP 145

RESULT 7

US-10-425-114-47346

; Sequence 47346, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 47346

; LENGTH: 584

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: 701043069_FLI.pep
US-10-425-114-47346

Query Match

Best Local Similarity 31.2%; Score 55.5; DB 15; Length 584;

Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 2 GQTLWGSQMEAWL-ATVLKALPWHPP 25

Db 136 EGTLMWHAHSSWLKATVYCALIHP 160

RESULT 8

US-10-437-963-124083

; Sequence 124083, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 124083

; LENGTH: 853

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_26857C.1.pep

US-10-437-963-124083

Query Match 30.9%; Score 55; DB 16; Length 853;

Best Local Similarity 37.0%; Pred. No. 1.7e+02;

Matches 10; Conservative 2; Mismatches 13; Indels 2; Gaps 1;

QY 1 GQTLWGSQMEAWL-ATVLKALPWHPP 27

Db 222 GNVTIWRSQQ--WTGQNFVGIWPRPLY 246

RESULT 9

US-09-995-494-102

; Sequence 102, Application US/09995494

; Patent No. US20020127578A1

; GENERAL INFORMATION:

; APPLICANT: Salceda, Susana

; APPLICANT: Macina, Roberto

; APPLICANT: Recipon, Herve

; APPLICANT: Cafferkey, Robert

; APPLICANT: Ali, Shujath

; APPLICANT: Sun, Yongming

; APPLICANT: Liu, Chenghua

; APPLICANT: Chen, Sei-Yu

; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0293

; CURRENT APPLICATION NUMBER: US/09/995,494

; CURRENT FILING DATE: 2001-11-27

; PRIOR APPLICATION NUMBER: 60/253,176

; NUMBER OF SEQ ID NOS: 115

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 102

; LENGTH: 48

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-995-494-102

Query Match 30.6%; Score 54.5; DB 9; Length 48;

Best Local Similarity 43.3%; Pred. No. 11;

Matches 13; Conservative 3; Mismatches 9; Indels 5; Gaps 2;

QY 1 GQTLW-GSGMEAWL-ATVLKALPWHPPYQL 29

Db 17 GIGLWWSGRGTN---TPVKSIGWHKTYQL 42

RESULT 10

US-10-425-115-235453

; Sequence 235453, Application US/10425115

; Publication No. US20040214272A1

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 235453
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_146314C.1.pep
US-10-425-115-235453

Query Match          30.6%; Score 54.5; DB 17; Length 242;
Best Local Similarity 40.0%; Pred. No. 56;
Matches 12; Conservative 4; Mismatches 13; Indels 1; Gaps 1;

QY 1 GGGTLWGSMEAWLATVLKALPWHPTTYOLE 30
      |||||
Db 183 GSGKVAWGTAACATYLRCKYHPTMEAE 211

RESULT 11
US-10-425-115-228082
; Sequence 228082, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 228082
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_139605C.1.pep
US-10-425-115-228082

Query Match          30.3%; Score 54; DB 17; Length 74;
Best Local Similarity 46.4%; Pred. No. 19;
Matches 13; Conservative 4; Mismatches 5; Indels 5; Gaps 2;

QY 6 WSGGMEAWLATVLKALPWHPTTYQL 29
      |||||
Db 23 WSGGVNGSTAGWPASLIALP--PPAQL 48

RESULT 12
US-10-282-122A-62303
; Sequence 62303, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62303
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62303

Query Match          30.3%; Score 54; DB 15; Length 575;
Best Local Similarity 38.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY 6 WSGGMEAWLATVLKALPWHPTTYQLEP 31
      |||||
Db 422 WNRVQAWLABETGQPPWTFWAGRP 447

RESULT 13
US-10-282-122A-64407
; Sequence 64407, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26

```

```
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 64407
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64407

Query Match      30.3%; Score 54; DB 15; Length 575;
Best Local Similarity 38.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY      6 WSGGMEAWLATVLKALPWHPTYQLEP 31
Db      422 WNRVQWLAEETGQPPWTPWYAGRP 447

RESULT 14
US-10-425-114-69836
; Sequence 69836, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69836
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROB73006B02_FLI.pep
US-10-425-114-69836

Query Match      30.1%; Score 53.5; DB 15; Length 542;
Best Local Similarity 52.0%; Pred. No. 1.7e+02;
Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY      2 OGTLWGSMEAWL-ATVLKALPWH 25
Db      54 EGTLWWHAHSSWLRAVHGALIHP 88

RESULT 15
US-10-425-115-198410
; Sequence 198410, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 198410
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(607)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_112527C.1.pep
US-10-425-115-198410

Query Match      30.1%; Score 53.5; DB 17; Length 607;
Best Local Similarity 52.0%; Pred. No. 1.9e+02;
Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY      2 OGTLWGSMEAWL-ATVLKALPWH 25
Db      129 EGTLWWHAHSSWLRAVHGALIHP 153

Search completed: November 11, 2004, 02:43:15
Job time : 36.9047 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 7.60377 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-59
Perfect score: 178

Sequence: 1 GQGLMGSGMEAWLATVLPWHTTYQLEP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.5	33.4	479	2 S68598	sucrose-6-phosphat
2	56.5	31.7	477	2 A33647	melittin resistanc
3	56	31.5	304	2 A10722	conserved hypothet
4	56	31.5	614	2 T09902	hypothetical prote
5	55	30.9	590	2 A86307	amino acid transpo
6	54	30.3	575	2 B70612	probable recD prot
7	54	30.3	576	2 S12792	protein-tyrosine k
8	54	30.3	888	2 I58378	tyrosine kinase -
9	53.5	30.1	454	2 A43501	sucrose-6-phosphat
10	53.5	30.1	762	2 A54411	amine oxidase (cop
11	51.5	28.9	559	2 T48180	hypothetical prote
12	51.5	28.9	763	2 JC5234	amine oxidase (cop
13	51	28.7	25	2 P70328	Ig heavy chain CDR
14	51	28.7	848	1 T02053	S receptor kinase
15	50	28.1	402	2 A18853	sodium-dependent n
16	50	28.1	464	2 A33625	protein-tyrosine k
17	50	28.1	480	2 T20899	hypothetical prote
18	50	28.1	577	2 T45333	hypothetical prote
19	50	28.1	753	2 G87178	probable DNA-bindi
20	50	28.1	864	2 A48266	protein-tyrosine k
21	49.5	27.8	286	2 B72404	Glycine-CRNA ligas
22	49.5	27.8	386	2 T14243	ubiquinol-cytochro
23	49	27.5	611	1 S06047	endo-1,4-beta-xyla
24	48.5	27.2	349	2 A52369	hypothetical prote
25	48.5	27.2	385	2 S53087	ubiquinol-cytochro
26	48.5	27.2	749	2 S77175	sensory transducti
27	48	27.0	88	2 I40663	9K protein-Coxi
28	48	27.0	179	2 A03484	phosphinothricin N
29	48	27.0	227	2 D84366	mannose-1-phosphat

30 48 27.0 248 2 T00300
31 48 27.0 248 2 JQ1340
32 48 27.0 294 2 AC9301
33 48 27.0 304 2 G85697
34 48 27.0 304 2 G85697
35 48 27.0 304 2 E64865
36 48 27.0 344 2 AH0041
37 48 27.0 374 2 C84339
38 48 27.0 475 2 AC1246
39 48 27.0 475 2 AG1608
40 48 27.0 503 2 E87101
41 48 27.0 614 2 S64447
42 48 27.0 622 2 S61140
43 48 27.0 831 2 S76235
44 47.5 26.7 142 2 T51902
45 47.5 26.7 251 1 ESDKTM

trax protein - Esc
trax protein - Esc
conserved hypothet
L, D-carboxyprptid
L, D-carboxyprptid
Muramoyltetrapepti
L-rhamnose-proton
quinolinate synthe
branched-chain alp
probable membrane
probable membrane
hypothetical prote
hypothetical prote
oleoyl-lacetyl-carri

ALIGNMENTS

RESULT 1

S68598 sucrose-6-phosphate hydrolase ScrB - Streptococcus sobrinus (strain 6715)
C;Species: Streptococcus sobrinus

A;Variety: Strain 6715
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 07-May-1999

C;Accession: S68598
R;Chen, Y.Y.M.; Lee, L.N.; LeBlanc, D.J.

Infect. Immun. 61, 2602-2610, 1993
A;Title: Sequence analysis of scrA and scrB from Streptococcus sobrinus 6715.

A;Reference number: S68598; MUID:93273516; PMID:8500898
A;Accession: S68598

A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-479 <CHE>
A;Cross-references: EMBL:L06792

C;Genetics:
A;Gene: scrB

C;Superfamily: sucrose-6-phosphate hydrolase

Query Match 33.4%; Score 59.5; DB 2; Length 479;
Best Local Similarity 39.3%; Pred. No. 4.3;
Matches 11; Conservative 3; Mismatches 7; Indels 7; Gaps 1;

Qy 11 EAWLATVLPKAL-----PWHPTYQLEP 31
Db 13 EDWSAEIKTIQENVAKSPWHTTYHIEP 40

RESULT 2

AB3647 melittin resistance protein pqaB [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AB3647

R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lete

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melite

A;Reference number: AD3252; PMID:11756688
A;Accession: AB3647

A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-477 <KUR>
A;Cross-references: UNIPROT:Q8YB02; GB:AE008918; PIDN:AAL54341.1; PID:gl7985322; GSPDB:

A;Experimental source: strain 16M.
C;Genetics:

A;Gene: BMEI11099
A;Map position: II

Query Match 31.7%; Score 56.5; DB 2; Length 477;
Best Local Similarity 39.3%; Pred. No. 11;

Matches 11; Conservative 4; Mismatches 8; Indels 5; Gaps 1;
 QY 3 GTLWGSME-----AWLATVVKALPWP 25
 Db 248 GDLYGAGREHARGSIWLFVATLWSP 275

RESULT 3
 AIO722
 conserved hypothetical protein STY1928 [imported] - Salmonella enterica subsp. enterica
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AIO722
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moulé, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AIO722
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-304 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD05483.1; PID:GL6502987; GSPDB:GN00176
 A:Genetics:
 A:Gene: STY1928

Query Match 31.5%; Score 56; DB 2; Length 304;
 Best Local Similarity 40.0%; Pred. No. 7.8;
 Matches 10; Conservative 4; Mismatches 9; Indels 2; Gaps 1;
 QY 2 QGTLWGSMEAWLATVVKALPWP 26
 Db 169 QGTLWGSNLAFLISLI--GTPWMP 191

RESULT 4
 T09902
 hypothetical protein T22A6.220 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T09902
 R:Bevan, M.; Zimmermann, W.; Gruenewald, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16896
 A:Accession: T09902
 A:Molecule type: DNA
 A:Residues: 1-614 <BEV>
 A:Cross-references: UNIPROT:O9STV5; EMBL:AL078637; GSPDB:GN00062; ATSP:T22A6.220
 A:Experimental source: cultivar Columbia; BAC clone T22A6
 C:Genetics:
 A:Gene: ATSP:T22A6.220
 A:Map position: 4
 A:Introns: 192/2; 358/3

Query Match 31.5%; Score 56; DB 2; Length 614;
 Best Local Similarity 44.4%; Pred. No. 16;
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 6 WGSMEAWLATVVKALPWP 23
 Db 125 WGAQFSPWVAATAKAYPW 142

RESULT 5
 A86307
 amino acid transporter homolog [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: A86307
 R:theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
 Rizzo, M.; Roney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: A86307
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-590 <STO>
 A:Cross-references: UNIPROT:Q9SHH0; GB:AE005172; NID:G5734765; PIDN:AAD50030.1; GSPDB:G
 C:Genetics:
 A:Map position: 1
 C:Superfamily: ecotropic retrovirus receptor protein

Query Match 30.9%; Score 55; DB 2; Length 590;
 Best Local Similarity 34.8%; Pred. No. 21;
 Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
 QY 1 GQGTWGSMEAWLATVVKALPW 23
 Db 482 GVSALWNSGVKGMIAVTVTGVW 504

RESULT 6
 B70612
 probable recD protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: B70612
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Accession: B70612
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-575 <COL>
 A:Cross-references: UNIPROT:P96919; GB:Z92772; GB:AL123456; NID:G3261722; PIDN:CAB07118
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: recD
 C:Superfamily: exodeoxyribonuclease V 67K chain

Query Match 30.3%; Score 54; DB 2; Length 575;
 Best Local Similarity 38.5%; Pred. No. 27;
 Matches 10; Conservative 2; Mismatches 14; Indels 0; Gaps 0;
 QY 6 WGSMEAWLATVVKALPWP 31
 Db 422 WNRVQAWLAETGQPPWTPWYGRP 447

RESULT 7
 S12792
 protein-tyrosine kinase (EC 2.7.1.112) ltk - mouse
 N:Alternate names: leukocyte tyrosine kinase
 C:Species: Mus musculus (house mouse)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-2004
 C:Accession: S12792; S00904
 R:Bernards, A.; de la Monte, S.M.
 EMBO J. 9, 2279-2287, 1990
 A:Title: The ltk receptor tyrosine kinase is expressed in pre-B lymphocytes and cerebra
 A:Reference number: S12792; MUID:90291994; PMID:2357970
 A:Accession: S12792
 A:Molecule type: mRNA

Ry:Rato, Y.; Poy, P.; Jacobson, G.R.; Kuramitsu, H.K.
J. Bacteriol. 171, 263-271, 1989
Title: Characterization and sequence analysis of the scrA gene encoding enzyme II (scrA)
Accession number: A32243; MUID:89123027; PMID:2536556

Query Match 30.1%; Score 53.5; DB 2; Length 454;
Best Local Similarity 34.6%; Pred. No. 25;
Matches 9; Conservative 6; Indels 7; Gaps 1;

QY 13 WLATVLKAL-----PWHTYOLEP 31
| : : | : : | : : | : : | : : |
15 WTBEIKSIKTNVALSFWHTTHIEP 40

RESULT 10

A:Accession: A54411
A:Molecule type: mRNA
A:Residues: 1-762 <MUA>
R:James, S.M.; Mu, D.; Wemmer, D.; Smith, A.J.; Brown, D.E.; Dooley, D.M.; Klinman, J.P.
J. Biol. Chem. 267, 7979-7982, 1992
A:Title: Tyrosine codon corresponds to topa quinone at the active site of copper amine
A:Reference number: A38081; MUID:92235001; PMID:1569055
A:Accession: B38081
A:Molecule type: protein
A:Residues: 463-465, 'D', 467-473, 'X', 475-485 <MU2>
R:James, S.M.; Mu, D.; Wemmer, D.; Smith, A.J.; Kaur, S.; Maltby, D.; Burlingame, A.L.; S
Science 248, 981-987, 1990
A:Title: A new redox cofactor in eukaryotic enzymes: 6-hydroxydopa at the active site c
A:Reference number: A48242; MUID:90260648; PMID:2111581
A:Accession: A48242
A>Status: preliminary
A:Molecule type: protein
A:Residues: 468-472 <JAN>
Ride Blase, D.; Acostinelli, E.; de Matteis, G.; Mondovi, B.; Morpurgo, L.
Eur. J. Biochem. 237, 93-99, 1996
A:Title: Half-of-the-sites reactivity of bovine serum amine oxidase. Reactivity and chr
A:Reference number: S65408; MUID:96203913; PMID:8620899
A:Accession: S65408
A:Molecule type: protein
A:Residues: 463-469, 'X', 471-487 <DEB>
C:Superfamily: amiloride-binding protein
C:Keywords: copper; glycoprotein; oxidoreductase; quinoicoprotein; topaquinone
F:1-16/Domains: signal sequence #status predicted <SIG>
F:17-762/Product: amine oxidase (copper-containing) #status predicted <MAT>
F:136,231,665/binding site: carbohydate (Asn) (covalent) #status predicted
F:470/Modified site: topaquinone (Tyr) #status experimental

Query Match 30.1%; Score 53.5; DB 2; Length 762;
Best Local Similarity 37.0%; Pred. No. 42;
Matches 10; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

QY 9 GNEAW-----LATVLKALPMWHTYOLE 30
| : : | : : | : : | : : | : : |
Db 533 GLENVWVAEDMAFVPITAPWSPEHQIQ 559

RESULT 11

T48180
hypothetical protein F7A7.110 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48180
R;Bevan, M.; Terry, N.; Ardiles, W.; Buyssehaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24487
A;Accession: T48180
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-559 <BEV>
A;Cross-references: UNIPROT:Q9M016; EMBL:AL161946
A;Experimental source: cultivar Columbia; BAC clone F7A7
C;Genetics:
A;Map position: 5
A;Intons: 232/1; 246/1; 268/2; 284/1; 339/1; 426/1; 478/3
A;Note: F7A7.110
C;Superfamily: Arabidopsis thaliana hypothetical protein F7A7.110

Query Match 28.9%; Score 51.5; DB 2; Length 559;
Best Local Similarity 37.0%; Pred. No. 55;
Matches 10; Conservative 2; Mismatches 8; Indels 7; Gaps 1;

QY 5 LMGSGMEAWLATVVKALPWHPTQLEP 31
DB 313 IWGEDMDWE-----APIHWVGLPEP 332

RESULT 12
JC5234
amine oxidase (copper-containing) (EC 1.4.3.6) AOC2 precursor - human
C;Species: Homo sapiens (man)
C;Date: 13-Mar-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
R;Accession: JC5234
R;Zhang, X.; McIntire, W.S.
Gene 179, 279-286, 1996
A;Title: Cloning and sequencing of a copper-containing, topa quinone-containing monoamin
A;Reference number: JC5234; MUID:97128319; PMID:8972912
A;Contents: placenta
A;Accession: JC5234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-763 <ZHA>
A;Cross-references: UNIPROT:Q16853; GB:U39447; NID:91399031; PID:913990
C;Comment: This enzyme catalyzes the oxidation of primary amines to the corresponding al
Cu(II) and 1 molecule of covalently-bound topa quinone. It also can scavenge circulating
C;Genetics:
A;Gene: GDB:AOC2; DAO2
A;Cross-references: GDB:4562632
A;Map position: 17q21-17q21
C;Superfamily: amiloride-binding protein
C;Keywords: copper; metalloprotein; oxidoreductase; quinoprotein; topaquinone
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-763/Product: amine oxidase (copper-containing) #status predicted <MAT>
F;444,520,522/Binding site: copper (His) #status predicted
F;471/Modified site: topaquinone (Tyr) #status predicted

Query Match 28.9%; Score 51.5; DB 2; Length 763;
Best Local Similarity 35.7%; Pred. No. 76;
Matches 10; Conservative 6; Mismatches 7; Indels 5; Gaps 1;

QY 8 SCWEAW----LATVVKALPWHPTQLE 30
DB 533 AGLNWNWAEDMVFVPMVAVPWSPEHQLQ 560

RESULT 13
PT0328
Ig heavy chain CDR3 region (clone J2-118) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 02-Feb-2001

C;Accession: PT0328
R;Yanada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A;Reference number: PT0222; MUID:91108337; PMID:1899102
C;Accession: PT0328
A;Molecule type: DNA
A;Residues: 1-25 <YAM>
A;Experimental source: B lymphocyte
A;Note: the authors translated the stop codon for residue 14 as X
C;Keywords: heterotetramer; immunoglobulin

Query Match 28.7%; Score 51; DB 2; Length 25;
Best Local Similarity 50.0%; Pred. No. 2.8;
Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 2;

QY 5 LMGSGMEAWLATVVKAL-PWHP 25
DB 5 LWFQGV-IMGXVLRLSLGPWHP 25

RESULT 14
T02053
S-receptor kinase (EC 2.7.1.-) KIK1 precursor - maize
N;Alternate names: K1 domain interacting kinase 1
C;Species: Zea mays (maize)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T02053
R;Braun, D.M.; Stone, J.M.; Walker, J.C.
submitted to the EMBL Data Library, December 1996
A;Description: Multiple receptor-like protein kinases interact with the K1 domain of KA
A;Reference number: Z14523
A;Accession: T02053
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-848 <BRA>
A;Cross-references: UNIPROT:O49974; EMBL:U82481; NID:92735016; PID: AAB93834.1; PID:927
A;Experimental source: strain B73
C;Genetics:
A;Gene: KIK1
C;Function:
A;Description: involved in preventing fertilization between plants having the same S-lo
C;Superfamily: S-receptor kinase; protein kinase homology; S-locus-specific glycoprotein
C;Keywords: ATP; glycoprotein; magnesium; phosphotransferase; receptor; serine/threonin
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-848/Product: S-receptor kinase KIK1 #status predicted <MAT>
F;33-435/Domain: S-locus-specific glycoprotein homology <SSG>
F;444-465/Domain: transmembrane #status predicted <TM>
F;528-808/Domain: protein kinase homology <KIN>
F;536-544/Region: protein kinase ATP-binding motif
F;111,123,218,263,274,340,356,393/Binding site: carbohydrate (Asn) (covalent) #status p
F;558,574,655,657/Active site: Lys, Glu, Asp, Lys #status predicted
F;660,664/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 28.7%; Score 51; DB 1; Length 848;
Best Local Similarity 35.7%; Pred. No. 98;
Matches 10; Conservative 2; Mismatches 14; Indels 2; Gaps 1;

QY 1 GQGLMGSGMEAWLATVVKALPWHPTQ 28
DB 217 GNVTYWRSQ--WNGVNFIGIFWRPLYR 242

RESULT 15
A11853
sodium-dependent nucleoside transporter all0378 [imported] - Nostoc sp. (strain PCC 712
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: A11853
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. PCC 7120
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: A11853
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-402 <KUR>
A;Cross-references: UNIPROT:Q8YVZ9; GB:BA000013; FIDN:BAF72336.1; PID:gl17129723; GSPDS:Q8YVZ9
A;Experimental source: strain PCC 7120
C;Genetics:
C;Gene: all0378
C;Superfamily: pyrimidine nucleoside transport protein nupC

A;Gene: all0378
C;Superfamily: pyrimidine nucleoside transport protein nupC

Query Match 28.1%; Score 50; DB 2; Length 402;
Best Local Similarity 50.0%; Pred. No. 62;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 6 WGSMEAWLATVLKALPW 23
|||:|:|:|:|:
Db 34 WGLGLEFVLALVILKTPW 51

Search completed: November 10, 2004, 14:52:26
Job time : 7.60377 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 41.2943 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-59

Perfect score: 178

Sequence: 1 GQGLWNGMEAWLTLKALPWHPTQLEP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
1	60.5	34.0	565	2 Q9FLB5	Q9flb5 arabidopsis
2	57	32.0	199	2 Q8RVJ4	Q8rvj4 pinus pinas
3	56.5	31.7	477	2 Q8YB02	Q8yb02 bruceella me
4	56.5	31.7	478	2 Q8FXE2	Q8fxe2 bruceella su
5	56	31.5	304	2 Q8Z689	Q8z689 salmonella
6	56	31.5	304	2 Q8ZP18	Q8zp18 salmonella
7	56	31.5	614	2 Q9STV5	Q9stv5 arabidopsis
8	56	31.5	623	2 Q8RWQ8	Q8rwq8 arabidopsis
9	55	30.9	305	2 Q6SL32	Q6sl32 armillifer
10	55	30.9	305	2 AAS00841	Aas00841 armillife
11	55	30.9	590	2 Q9SHH0	Q9shh0 arabidopsis
12	55	30.9	599	2 Q7WMJ5	Q7wmj5 porphyromon
13	55	30.9	853	2 Q7EZ32	Q7ez32 oryza sativ
14	54.5	30.6	952	2 Q6Q0N0	Q6q0n0 rattus norv
15	54.5	30.6	952	2 AAS75317	Aas75317 rattus no
16	54	30.3	307	2 Q7SC19	Q7sc19 neurospora
17	54	30.3	561	2 Q7D914	Q7d914 mycobacteri
18	54	30.3	561	2 P96919	P96919 mycobacteri
19	54	30.3	575	2 Q7U1L1	Q7u1l1 mycobacteri
20	54	30.3	707	2 Q98JY6	Q98jy6 rhizobium l
21	54	30.3	888	1 KLTK_MOUSE	P08923 mus musculu
22	53.5	30.1	418	2 Q7JCD2	Q7jcd2 bacillus ce
23	53.5	30.1	418	2 AAS40064	Aas40064 bacillus
24	53.5	30.1	479	1 SCRB_STRMU	P13522 streptococc
25	53.5	30.1	762	1 AOXC_BOVIN	Q29437 bos taurus
26	53	29.8	341	2 Q7UWV6	Q7uwv6 rhodospirell
27	53	29.8	412	2 Q8VED1	Q8ved1 mus musculu
28	53	29.8	439	2 Q8FFJ4	Q8ffj4 arabidopsis
29	52.5	29.5	240	2 Q9DZ20	Q9dzz0 mus musculu
30	52.5	29.5	553	2 Q87IG3	Q87ig3 vibrio para
31	52.5	29.5	734	1 PSAB_CYACA	Q9tlq6 cyanidium c

32	52.5	29.5	944	1 VGLB_TUHV2	Q9wrl5 tupaiid herp
33	52.5	29.5	944	2 Q77L59	Q77l59 tupaiia herp
34	52.5	29.5	944	2 AAK57104	Aak57104 tupaiia he
35	52	29.2	166	2 Q6XCA0	Q6xca0 mustelus ma
36	52	29.2	166	2 AAP04340	Aap04340 mustelus
37	52	29.2	227	2 Q8HMB2	Q8hmb2 lamprogramm
38	52	29.2	363	2 Q9QTE4	Q9qte4 marek's dis
39	52	29.2	437	2 Q9E942	Q9e942 gallid herp
40	52	29.2	468	2 Q7UN7	Q7un7 synchococc
41	52	29.2	471	2 Q7UR92	Q7ur92 rhodospirell
42	52	29.2	883	2 Q9RF11	Q9rf11 myxococcus
43	51.5	28.9	238	2 Q76E57	Q76e57 homo sapien
44	51.5	28.9	238	2 BAD12557	Bad12557 homo sapi
45	51.5	28.9	240	2 Q804J1	Q804j1 brachydanio

ALIGNMENTS

RESULT 1					
Q9FLB5	Q9FLB5	PRELIMINARY;	PRT;	565 AA.	
AC	Q9FLB5	PRELIMINARY;	PRT;	565 AA.	
DT	01-MAR-2001	(TREMBlrel. 16, Created)			
DT	01-MAR-2001	(TREMBlrel. 16, Last sequence update)			
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)			
DE	Laccase (Diphenol oxidase).				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.				
OX	NCBI_TaxID=3702;				
[1]					
RN	SEQUENCE FROM N.A.				
RP	MEDLINE=98344145; PubMed=9679202;				
RX	Caneko T.; Kotani H.; Nakamura Y.; Sato S.; Asamizu E.; Miyajima N.;				
RA	Tabata S.;				
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence				
RT	features of the regions of 1,381,565 bp covered by twenty one				
RT	physically assigned P1 and TAC clones."				
RL	DNA Res. 5:131-145(1998).				
DR	EXBL; AB010692; BAB09982.1; -.				
DR	HSP; Q9Y780; IHFU.				
DR	GO; GO:0005507; F: copper ion binding; IEA.				
DR	InterPro; IPR001117; Cu-oxidase.				
DR	InterPro; IPR008972; Cupredoxin.				
DR	InterPro; IPR002355; Cu ox copper_BS.				
DR	Pfam; PF00394; Cu-oxidase_1.				
DR	PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.				
DR	PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.				
SQ	SEQUENCE 565 AA; 62733 MW; 15DA9DD5CA8932E CRC64;				

Query Match 34.0%; Score 60.5; DB 2; Length 565;
Best Local Similarity 53.8%; Pred. No. 23;
Matches 14; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

Qy	2 QGLWNGMEAWL-ATVLKALPWHPT 26
Db	121 EGTLWHAHSHSLRATVYGALIIHPT 146

RESULT 2					
Q8RVJ4	Q8RVJ4	PRELIMINARY;	PRT;	199 AA.	
AC	Q8RVJ4	PRELIMINARY;	PRT;	199 AA.	
DT	01-JUN-2002	(TREMBlrel. 21, Created)			
DT	01-JUN-2002	(TREMBlrel. 21, Last sequence update)			
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)			
DE	Putative acid phosphatase (Fragment).				
OS	Pinus pinaster (Maritime pine)				
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;				
OC	Spermatophyta; Coniferopsida; Pinaceae; Pinus.				
OX	NCBI_TaxID=71647;				

Dolichyl-phosphate-mannose-protein mannosyltransferase family protein.

DE DE STRAIN=1330 / Biovar 1;
GN OrderedLocusNames=BRAO137;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OX Brucellaceae; Brucella.
NCBI_TaxID=29461;
[1]

RN
RN SEQUENCE FROM N.A.
RP
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=2247741; PubMed=12771122; DOI=10.1073/pnas.192319099;

DR Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.P., Madupu R.,
RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tetzelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Hailing S.M., Boyle S.M., Fraser C.M.;

RT "The Brucella suis genome reveals fundamental similarities between animal and plant pathogens and symbionts.";
RL Proc. Natl Acad. Sci U.S.A. 99:13148-13153(2002).
EMBL; AB014517; AA033345.1; - .
DR TIGR; BRAO137; - .
DR GO; GO:0016020; C:membrane; IEA.
DR DR GO; GO:0000030; F:mannosyltransferase activity; IEA.
DR DR GO; GO:0016757; F:transferase activity, transferring glycosyl. ; IEA.
DR GO; GO:0006493; P-O-linked glycosylation; IEA.
DR InterPro; IPRO03342; Glyco_trans_39.
DR Pfam; PF02366; PMT; 1.
SQ Complete proteome; Glycosyltransferase; Transferase.
KW SEQUENCE 478 AA; 51957 MW; AB668D36AFECBCE1 CRC64;

Query Match 31.7%; Score 56.5; DB 2; Length 478;
Best Local Similarity 39.3%; Pred. No. 65;
Matches 11; Conservative 4; Mismatches 8; Indels 5; Gaps 1;

OY 3 GTLMGSGME-----AWLATVLKALPWWP 25
||:||| || | : ||:
DB 249 GDLYAGREHARGSIWLFIVATLPWSP 276

RESULT 5
Q8Z689 PRELIMINARY; PRT; 304 AA.
ID Q8Z689 AC Q8Z689; Q7CAF6;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE Hypothetical protein STV1928.
DN OrderedLocusNames-STV1928.t1077;
GN Salmonella typhi.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=Ct18;
RX MEDLINE=N21534947; PubMed=11677608; DOI=10.1038/35101607;

DR Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mangall K.I., Bentley S.D., Holden M.T.G., Sebaihia M., Baker S., Basham D., Brooks K., Chillingworth T., Connerton P., Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J., Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K., Krogh A., Larsen I.S., Leatherbarrow S., Moule S., O'Garra P., Parry C., Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K., Whithead S., Barrrell B.G.;

RA "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typh Ct18."
RT Nature 413:848-852(2001).

RL [2]
RN
RN SEQUENCE FROM N.A.
RP
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;

RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.,
RT "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2
and CT18";
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627272; CAD05483.1; -;
DR MEROPS; U61.001; -;
DR InterPro; IPR003507; Peptidase U61.
DR Pfam; PF02016; Peptidase U61; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 304 AA; 33355 MW; 26D9EA86D2DEC88E CRC64;
Query Match 31.5%; Score 56; DB 2; Length 304;
Best Local Similarity 40.0%; Pred. No. 49;
Matches 10; Conservative 4; Mismatches 9; Indels 2; Gaps 1;
Qy 2 QGTLWGSMEAWLATVVKALPWHPT 26
Db 169 QGTLWGSMEAWLATVVKALPWHPT 191
RESULT 6
Q8ZP18 PRELIMINARY; PRT; 304 AA.
AC Q8ZP18
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative resistance protein MccP.
GN Name=ygq; OrderedLocusNames=STM1800;
OS *Salmonella typhimurium*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen K., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea R.K., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
LT2";
RL Nature 413:852-856(2001).
DR EMBL; AE008780; AAL20715.1; -;
DR MEROPS; U61.001; -;
DR InterPro; IPR003507; Peptidase U61.
DR Pfam; PF02016; Peptidase U61; 1.
KW Complete proteome.
SQ SEQUENCE 304 AA; 33351 MW; 2CEB67DDC06DD6D5 CRC64;
Query Match 31.5%; Score 56; DB 2; Length 304;
Best Local Similarity 40.0%; Pred. No. 49;
Matches 10; Conservative 4; Mismatches 9; Indels 2; Gaps 1;
Qy 2 QGTLWGSMEAWLATVVKALPWHPT 26
Db 169 QGTLWGSMEAWLATVVKALPWHPT 191
RESULT 7
Q9STV5 PRELIMINARY; PRT; 614 AA.
AC Q9STV5
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Transport inhibitor response-like protein.
GN Name=T22A6.220; Synonyms=AT4g24390;
OS *Arabidopsis thaliana* (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Bancroft I.,
RA Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Zimmermann W., Gruenisen A., Wambutt R., Kalicki J., Wohldmann P.,
RA Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL078637; CAB45074.1; -;
DR EMBL; AL161561; CAB79349.1; -;
DR PIR; T09902; T09902.
DR InterPro; IPR001810; P-box.
DR Pfam; PF00646; P-box; 1.
SQ SEQUENCE 614 AA; 68520 MW; A856C01DF54ADB9 CRC64;
Query Match 31.5%; Score 56; DB 2; Length 614;
Best Local Similarity 44.4%; Pred. No. 96;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
Qy 6 WGSMEAWLATVVKALPWHPT 23
Db 125 WGAQFSPWVAATAKAVFW 142
RESULT 8
Q8RWQ8 PRELIMINARY; PRT; 623 AA.
AC Q8RWQ8
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AT4g24390/T22A6.220.
OS *Arabidopsis thaliana* (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Huan V.W., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY091772; AAM10320.1; -;
DR EMBL; BT006340; AAP21148.1; -;

[illegible]

```
Query Match      30.9%; Score 55; DB 2; Length 590;
Best Local Similarity 34.8%; Pred. No. 1.2e+02;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY      1 GQTLWGSMEAWLATVLKALPW 23
DB      482 GVSALWNSGVKGIATVTGVIV 504

RESULT 12
Q7MWJ5 PRELIMINARY; PRT; 599 AA.
AC Q7MWJ5;
DT 01-VAR-2004 (TREMBLrel. 26, Created)
DT 01-VAR-2004 (TREMBLrel. 26, Last sequence update)
DE GTP-binding protein TypA.
GN Name=typA; OrderedLocNames=P00615;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.B., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83."
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AE017174; AAQ65793.1; -.
DR TIGR; PG0615; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0003746; P:translation elongation factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR000640; EFG C.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; ProtSyn_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR006298; trPA.
DR Pfam; PF00679; EFG C; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR TIGRFAMs; TIGR01394; TypA_BipA; 1.
DR PROSITE; PS00301; EFACTOR_GTP; 1.
KW Complete proteome; GTP-binding.
SQ SEQUENCE 599 AA; 67195 MW; A3FC5B93046824E3 CRC64;

Query Match      30.9%; Score 55; DB 2; Length 599;
Best Local Similarity 35.7%; Pred. No. 1.3e+02;
Matches 10; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

QY      4 TLWGSMEAWLATVLKALPWHPTQLEP 31
DB      162 TIYSAKQGWSTDYK---HPTQDIAP 185

RESULT 13
Q7EZ32 PRELIMINARY; PRT; 853 AA.
AC Q7EZ32;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative S-receptor kinase KIK1.

Name=OSUNB0041B22.1.09;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AP005126; BAC65367.1; -.
DR GO; GO:0016301; F:Kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR001480; E lectin.
DR InterPro; IPR000742; EGF_L2.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR000858; Slocus_glycop.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF01453; B lectin; 1.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; prot_kinase; 1.
DR SMART; SM00108; B lectin; 1.
DR SMART; SM00473; PAN_APP; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00927; BULB_LECTIN; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS50948; PAN; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferrase.
SQ SEQUENCE 853 AA; 94610 MW; 2294AD9F81CF29B CRC64;

Query Match      30.9%; Score 55; DB 2; Length 853;
Best Local Similarity 37.0%; Pred. No. 1.8e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 2; Gaps 1;

QY      1 GQTLWGSMEAWLATVLKALPWHPTY 27
DB      222 GNVTVRSQ--WTGQNFVGIPWRPLY 246

RESULT 14
Q6Q0N0 PRELIMINARY; PRT; 952 AA.
AC Q6Q0N0;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Calsynthenin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RA Ahmed F., Torrado M., Zinovieva R.D., Senatorov V., Wistow G.,
RA Tomarev S.I.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 cadherin domains.
DR EMBL; AY569014; AS575317.1; -.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR Pfam; PF00028; Cadherin; 2.
DR PRINTS; PR00205; CADHERIN.
```

DR SMART; SM00112; CA; 2.
 DR PROSITE; PS0268; CADHERIN_2; 2.
 KW Calcium; Calcium-binding.
 SQ SEQUENCE 952 AA; 106260 MW; DBE2185AF20F9713 CRC64;
 Query Match 30.6%; Score 54.5; DB 2; Length 952;
 Best Local Similarity 50.0%; Pred. No. 2.3e+02;
 Matches 13; Conservative 2; Mismatches 8; Indels 3; Gaps 2;
 QY 3 GTLWGSMEAWLATVLKALPW-HPTY 27
 DB 19 GLLYGGGV--WAARVNKHKFWLEPTY 42
 RESULT 15
 AAS75317
 ID AAS75317 PRELIMINARY; PRT; 952 AA.
 AC AAS75317;
 DT 31-MAR-2004 (TrEMBLrel. 27, Created)
 DT 31-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 31-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Calsynthenin.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar;
 RA Ahmed F., Torrado M., Zinovieva R.D., Senatorov V., Wistow G.,
 RA Tomarev S.I.;
 RT "Gene Expression Profile of the Rat Eye Irido-Corneal Angle. NEIBank
 RT Expressed Sequence Tag Analysis."
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY569014; AAS75317.1; -;
 SQ SEQUENCE 952 AA; 106260 MW; DBE2185AF20F9713 CRC64;
 Query Match 30.6%; Score 54.5; DB 2; Length 952;
 Best Local Similarity 50.0%; Pred. No. 2.3e+02;
 Matches 13; Conservative 2; Mismatches 8; Indels 3; Gaps 2;
 QY 3 GTLWGSMEAWLATVLKALPW-HPTY 27
 DB 19 GLLYGGGV--WAARVNKHKFWLEPTY 42
 Search completed: November 10, 2004, 14:50:29
 Job time : 44.2943 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 10.2491 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-60

Perfect score: 144

Sequence: 1 IAQATKATIDKWCNKIKIFYTSKKEAS 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS-COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	50.7	60	4	US-09-513-999C-6532
2	71	49.3	107	4	US-09-513-999C-8172
3	48	33.3	43	4	US-09-270-767-37159
4	48	33.3	43	4	US-09-270-767-52376
5	47	32.6	732	4	US-09-270-767-38665
6	47	32.6	732	4	US-09-270-767-53882
7	46	31.9	278	4	US-09-270-767-39161
8	46	31.9	278	4	US-09-270-767-54378
9	45.5	31.6	134	4	US-09-107-532A-5878
10	45	31.2	198	4	US-09-134-000C-5033
11	45	31.2	250	4	US-09-270-767-52113
12	45	31.2	381	4	US-09-270-767-43720
13	45	31.2	650	4	US-09-270-767-42511
14	44	30.6	60	4	US-09-328-352-6326
15	44	30.6	476	4	US-09-248-796A-16327
16	43.5	30.2	190	1	US-08-106-981-2
17	43	29.9	119	4	US-09-107-532A-6608
18	42.5	29.5	159	4	US-09-270-767-34992
19	42.5	29.5	159	4	US-09-270-767-50209
20	42.5	29.5	177	4	US-09-270-767-37349
21	42.5	29.5	177	4	US-09-270-767-52566
22	42	29.2	192	4	US-09-319-588C-8
23	42	29.2	193	4	US-09-248-796A-15614
24	42	29.2	535	4	US-09-248-796A-18860
25	42	29.2	1024	4	US-09-562-737-41
26	42	29.2	2710	1	US-08-480-604A-6
27	42	29.2	2710	2	US-08-405-496A-6

```

28 42 29.2 2710 3 US-08-915-136-6 Sequence 6, Appli
29 42 29.2 2710 3 US-08-957-310-6 Sequence 6, Appli
30 42 29.2 2710 4 US-10-011-366-6 Sequence 6, Appli
31 42 29.2 2710 4 US-09-084-517-6 Sequence 6, Appli
32 41.5 28.8 286 4 US-09-134-000C-4791 Sequence 4791, Ap
33 41 28.5 37 4 US-03-270-767-61460 Sequence 61460, A
34 41 28.5 127 2 US-08-574-699A-2 Sequence 2, Appli
35 41 28.5 194 4 US-09-270-767-45922 Sequence 45922, A
36 41 28.5 198 4 US-09-270-767-32467 Sequence 32467, A
37 41 28.5 198 4 US-09-270-767-47684 Sequence 47684, A
38 41 28.5 265 4 US-09-248-796A-15796 Sequence 15796, A
39 41 28.5 268 4 US-09-371-338-21 Sequence 21, Appli
40 41 28.5 311 4 US-09-912-628-6 Sequence 6, Appli
41 41 28.5 494 4 US-09-252-991A-24183 Sequence 24183, A
42 41 28.5 690 4 US-09-371-338-19 Sequence 19, Appli
43 41 28.5 724 4 US-09-270-767-44066 Sequence 44066, A
44 41 28.5 1024 4 US-09-562-737-46 Sequence 46, Appli
45 41 28.5 1294 2 US-08-819-288-3 Sequence 3, Appli

```

ALIGNMENTS

RESULT 1
US-09-513-999C-6532
; Sequence 6532, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6532
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 49
; OTHER INFORMATION: Xaa-Ile or Lys
US-09-513-999C-6532

Query Match 50.7%; Score 73; DB 4; Length 60;
Best Local Similarity 62.5%; Pred.No. 0.00018;
Matches 15; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 AQATKATIDKWCNKIKIFYTSKK 25
Db 12 AMATKIDKWDLIKLSFCTAKE 35

RESULT 2
US-09-513-999C-8172
; Sequence 8172, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487

us-10-092-750-60.ral

Fri Nov 12 14:55:47 2004

```

; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 8172
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-8172

Query Match      49.3%; Score 71; DB 4; Length 107;
Best Local Similarity 62.5%; Pred. No. 0.00069;
Matches 15; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy  2 AQATKATIDKNCIKLKIFYTSKK 25
Db  12 AWATKAKIDKWDLIKLSCTAKE 35

RESULT 3
US-09-270-767-37159
; Sequence 37159, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37159
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-37159

Query Match      33.3%; Score 48; DB 4; Length 43;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy  8 TIDKNCIKLKIFYTS 23
Db  1 TSDCYNVRLFIYSS 16

RESULT 4
US-09-270-767-52376
; Sequence 52376, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52376
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-52376

Query Match      33.3%; Score 48; DB 4; Length 43;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy  8 TIDKNCIKLKIFYTS 23
Db  1 TSDCYNVRLFIYSS 16

RESULT 5
US-09-270-767-52376
; Sequence 52376, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52376
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-52376

Query Match      33.3%; Score 48; DB 4; Length 43;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy  8 TIDKNCIKLKIFYTS 23
Db  1 TSDCYNVRLFIYSS 16

US-09-270-767-38665
; Sequence 38665, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38665
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-38665

Query Match      32.6%; Score 47; DB 4; Length 732;
Best Local Similarity 69.2%; Pred. No. 39;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy  13 NCIKLKIFYTSKK 25
Db  588 NCIKLLLYRSIK 600

RESULT 6
US-09-270-767-53882
; Sequence 53882, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53882
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-53882

Query Match      32.6%; Score 47; DB 4; Length 732;
Best Local Similarity 69.2%; Pred. No. 39;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy  13 NCIKLKIFYTSKK 25
Db  588 NCIKLLLYRSIK 600

RESULT 7
US-09-270-767-39161
; Sequence 39161, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39161
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-39161

```

```

;
; FEATURE:
;   OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39161

Query Match          31.9%;   Score 46;   DB 4;   Length 278;
Best Local Similarity 40.9%;   Pred. No. 20;
Matches          9;   Conservative          5;   Mismatches          8;   Indels          0;   Gaps          0;

QY      3   QATKATIDKWNCKIKLIFYTSK 24
      ||||| :|: :|: |||
Db      188  QATPLTICASCLDFRLYITSK 209

RESULT 8
US-09-270-767-54378
; Sequence 54378, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;   APPLICANT: Homburger et al.
;   TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;   FILE REFERENCE: File Reference: 7326-094
;   CURRENT APPLICATION NUMBER: US/09/270,767
;   CURRENT FILING DATE: 1999-03-17
;   NUMBER OF SEQ ID NOS: 62517
;   SOFTWARE: Patent In Ver. 2.0
;   SEQ ID NO 54378
;   LENGTH: 278
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54378

Query Match          31.9%;   Score 46;   DB 4;   Length 278;
Best Local Similarity 40.9%;   Pred. No. 20;
Matches          9;   Conservative          5;   Mismatches          8;   Indels          0;   Gaps          0;

QY      3   QATKATIDKWNCKIKLIFYTSK 24
      ||||| :|: :|: |||
Db      188  QATPLTICASCLDFRLYITSK 209

RESULT 9
US-09-107-532A-5878
; Sequence 5878, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
;   APPLICANT: Lynn A. Doucette-Stamm and David Bush
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                                     ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;   NUMBER OF SEQUENCES: 7310
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;   STREET: 100 Beaver Street
;   CITY: Waltham
;   STATE: Massachusetts
;   COUNTRY: USA
;   ZIP: 02354
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: CD-ROM ISO9660
;   COMPUTER: PC
;   OPERATING SYSTEM: <Unknown>
;   SOFTWARE: ASCII
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/107,532A
;   FILING DATE: 30-Jun-1998
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/085,598
;   FILING DATE: 14 May 1998
;   APPLICATION NUMBER: 60/051571
;   FILING DATE: July 2, 1997
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Ariniello, Pamela Deneke

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 32.3849 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-60

Perfect score: 144

Sequence: 1 IAQATKATIDKNCIKIKIFYTSKKEAS 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	144	100.0	28	14	US-10-092-750-60
2	79	54.9	317	13	US-10-083-853-2
3	79	54.9	464	15	US-10-211-462-221
4	79	54.9	641	16	US-10-408-765A-181
5	79	54.9	843	15	US-10-001-885-115
6	79	54.9	917	13	US-10-001-835-189
7	79	54.9	940	14	US-10-082-830-235
8	79	54.9	1010	14	US-10-050-882-80
9	79	54.9	1031	14	US-10-082-828A-257
10	79	54.9	1081	15	US-10-085-198-150
11	79	54.9	1272	15	US-10-114-270-30
12	79	54.9	1274	16	US-10-415-615-3
13	78	54.2	219	14	US-10-094-749-1795

```

14 53.5 101 9 US-09-867-550-1160 Sequence 1160, Ap
15 52.8 647 14 US-10-021-660-94 Sequence 94, Appl
16 75 52.1 990 15 US-10-085-198-154 Sequence 154, App
17 73 50.7 1018 15 US-10-085-198-156 Sequence 156, App
18 73 50.7 1275 14 US-10-025-201-3 Sequence 3, Appli
19 72 50.0 126 15 US-10-108-260A-4667 Sequence 4667, Ap
20 72 50.0 956 14 US-10-000-256A-153 Sequence 153, App
21 71 49.3 123 14 US-10-104-047-3008 Sequence 2008, Ap
22 71 49.3 206 14 US-10-104-047-3744 Sequence 3744, Ap
23 70 48.6 146 15 US-10-108-260A-2729 Sequence 2729, Ap
24 67 46.5 1017 15 US-10-085-198-152 Sequence 152, App
25 54 37.5 48 9 US-09-925-300-1097 Sequence 1097, Ap
26 54 37.5 81 15 US-10-276-774-1716 Sequence 1716, Ap
27 51 35.4 103 15 US-10-276-774-2492 Sequence 2492, Ap
28 50 34.7 120 15 US-10-424-599-178583 Sequence 178583,
29 49 34.0 1013 15 US-10-092-900A-230 Sequence 230, App
30 49 34.0 1013 15 US-10-336-472-128 Sequence 128, App
31 49 34.0 1013 15 US-10-236-472-152 Sequence 152, App
32 49 34.0 1018 15 US-10-336-472-126 Sequence 126, App
33 49 34.0 1018 15 US-10-236-472-154 Sequence 154, App
34 48 33.3 49 15 US-10-424-599-188449 Sequence 188449,
35 46.5 32.3 111 16 US-10-767-701-44455 Sequence 44455, A
36 46.5 32.3 186 17 US-10-425-115-280905 Sequence 280905,
37 46 31.9 87 17 US-10-425-115-233963 Sequence 233963,
38 46 31.9 373 16 US-10-479-284-30 Sequence 30, Appl
39 46 31.9 1842 14 US-10-243-552-847 Sequence 847, App
40 46 31.9 1842 14 US-10-243-552-887 Sequence 887, App
41 45.5 31.6 94 15 US-10-335-977-8018 Sequence 8018, Ap
42 45.5 31.6 127 15 US-10-335-977-8019 Sequence 8019, Ap
43 45.5 31.6 497 9 US-09-981-353-105 Sequence 105, App
44 45.5 31.6 508 14 US-10-170-385-347 Sequence 347, App
45 45 31.2 110 17 US-10-425-115-246081 Sequence 246081,

```

ALIGNMENTS

RESULT 1

```

US-10-092-750-60
; Sequence 60, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-60

```

```

Query Match 100.0%; Score 144; DB 14; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 IAQATKATIDKNCIKIKIFYTSKKEAS 28

Db 1 IAQATKATIDKNCIKIKIFYTSKKEAS 28

RESULT 2

```

US-10-083-853-2
; Sequence 2, Application US/10083853
; Publication No. US20020164709A1
; GENERAL INFORMATION:

```

; APPLICANT: Affymetrix, Inc
 ; APPLICANT: Shigetani, Ron T
 ; APPLICANT: Siani-Rose, Michael A
 ; TITLE OF INVENTION: Nucleic Acid Encoding Growth Factor Protein
 ; FILE REFERENCE: 3385.1
 ; CURRENT APPLICATION NUMBER: US/10/083,853
 ; PRIOR FILING DATE: 2002-02-26
 ; PRIOR APPLICATION NUMBER: USN 60/272,663
 ; PRIOR FILING DATE: 2001-03-01
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 317
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; ORGANISM: Homo Sapiens
 US-10-083-853-2

Query Match 54.9%; Score 79; DB 13; Length 317;
 Best Local Similarity 66.7%; Pred. No. 0.0014;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATATIDKNCIKIKIFYTSKK 25
 DB 12 AMATKAKIDKNDLIKKSFTAKE 35

RESULT 3
 US-10-211-462-221
 ; Sequence 221, Application US/10211462
 ; Publication No. US2004003495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
 ; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
 ; FILE REFERENCE: 018501-006200US
 ; CURRENT APPLICATION NUMBER: US/10/211,462
 ; CURRENT FILING DATE: 2003-02-13
 ; PRIOR APPLICATION NUMBER: US 09/784,356
 ; PRIOR FILING DATE: 2001-02-14
 ; PRIOR APPLICATION NUMBER: US 09/791,390
 ; PRIOR FILING DATE: 2001-02-22
 ; PRIOR APPLICATION NUMBER: US 60/310,025
 ; PRIOR FILING DATE: 2001-08-03
 ; PRIOR APPLICATION NUMBER: US 60/334,244
 ; PRIOR FILING DATE: 2001-11-29
 ; NUMBER OF SEQ ID NOS: 230
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 221
 ; LENGTH: 464
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; ORGANISM: Homo sapiens
 US-10-211-462-221

Query Match 54.9%; Score 79; DB 15; Length 464;
 Best Local Similarity 66.7%; Pred. No. 0.0021;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATATIDKNCIKIKIFYTSKK 25
 DB 12 AMATKAKIDKNDLIKKSFTAKE 35

RESULT 4
 US-10-408-765A-181
 ; Sequence 181, Application US/10408765A
 ; Publication No. US2004010187A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 ; FILE REFERENCE: 650088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 181
 ; LENGTH: 641
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; ORGANISM: Homo sapiens
 US-10-408-765A-181

Query Match 54.9%; Score 79; DB 16; Length 641;
 Best Local Similarity 66.7%; Pred. No. 0.0023;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATATIDKNCIKIKIFYTSKK 25
 DB 367 AMATKAKIDKNDLIKKSFTAKE 390

RESULT 5
 US-10-001-885-115
 ; Sequence 115, Application US/10001885
 ; Publication No. US20040058319A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Salceda, Susana
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Recipon, Herve
 ; APPLICANT: Cafferkey, Robert
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Liu, Chenguang
 ; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pr
 ; FILE REFERENCE: DEX-0279
 ; CURRENT APPLICATION NUMBER: US/10/001,885
 ; CURRENT FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 60/252,061
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: 60/253,257
 ; PRIOR FILING DATE: 2000-11-27
 ; NUMBER OF SEQ ID NOS: 167
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 115
 ; LENGTH: 843
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; ORGANISM: Homo sapien
 US-10-001-885-115

Query Match 54.9%; Score 79; DB 15; Length 843;
 Best Local Similarity 66.7%; Pred. No. 0.0039;
 Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATATIDKNCIKIKIFYTSKK 25
 DB 715 AMATKAKIDKNDLIKKSFTAKE 738

RESULT 6
 US-10-001-835-189
 ; Sequence 189, Application US/10001835
 ; Publication No. US20020160387A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Salceda, Susana
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Recipon, Herve
 ; APPLICANT: Cafferkey, Robert
 ; APPLICANT: Sun, Yongming

```
; APPLICANT: Liu, Chenghua
; FILE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; TITLE REFERENCE: DEX-0277
; CURRENT APPLICATION NUMBER: US/10/001,835
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,997
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 189
; TYPE: PRT
; LENGTH: 917
; ORGANISM: Homo sapien
US-10-001-835-189

Query Match      54.9%; Score 79; DB 13; Length 917;
Best Local Similarity 66.7%; Pred. No. 0.0043;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 AOAATKATIDKWNCKIKLIFYTSKK 25
DB      790 AATKAKIDKNDLKLKSFCTAKE 813

RESULT 7
US-10-082-830-235
; Sequence 235, Application US/10082830
; Publication No. US20030077604A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0249
; CURRENT APPLICATION NUMBER: US/10/082,830
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,802
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 235
; TYPE: PRT
; LENGTH: 940
; ORGANISM: Homo sapiens
US-10-082-830-235

Query Match      54.9%; Score 79; DB 14; Length 940;
Best Local Similarity 66.7%; Pred. No. 0.0044;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 AOAATKATIDKWNCKIKLIFYTSKK 25
DB      812 AATKAKIDKNDLKLKSFCTAKE 835

RESULT 8
US-10-050-882-80
; Sequence 80, Application US/10050882
; Publication No. US20030104400A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 27 Human secreted proteins
; FILE REFERENCE: P2038P1
; CURRENT APPLICATION NUMBER: US/10/050,882
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/661,453
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: PCT/US00/06783
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/125,055

; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
; LENGTH: 1010
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (104)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (194)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (362)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (525)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (643)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (649)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (656)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (660)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (731)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (770)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (777)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (790)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (800)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (825)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (987)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (996)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (1003)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-050-882-80

Query Match      54.9%; Score 79; DB 14; Length 1010;
Best Local Similarity 66.7%; Pred. No. 0.0047;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 AOAATKATIDKWNCKIKLIFYTSKK 25
DB      736 AATKAKIDKNDLKLKSFCTAKE 759
```

```

RESULT 9
US-10-082-828A-257
; Sequence 257, Application US/10082828A
; Publication No. US2003017515A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and P
; FILE REFERENCE: DEX-0247
; CURRENT APPLICATION NUMBER: US/10/082,828A
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US 60/243,805
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 257
; LENGTH: 1031
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-828A-257

Query Match          54.9%; Score 79; DB 14; Length 1031;
Best Local Similarity 66.7%; Pred. No. 0.0048;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY  2 A0ATKATIDKWCICIKLIFYTSKK 25
    |||||:||||:||||:||||:
DB  903 AMATKAKIDKWDLIKLSFCTAKE 926

RESULT 10
US-10-085-198-150
; Sequence 150, Application US/10085198
; Publication No. US2004000907A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-279
; CURRENT APPLICATION NUMBER: US/10/085,198
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/271,646
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/276,401
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,981
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,858
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/271,840
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/277,324
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/286,096
; PRIOR FILING DATE: 2001-04-21
; PRIOR APPLICATION NUMBER: 60/299,695
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/315,614
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/272,405
; PRIOR FILING DATE: 2001-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 653
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 150
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-085-198-150

Query Match          54.9%; Score 79; DB 15; Length 1081;
Best Local Similarity 66.7%; Pred. No. 0.005;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY  2 A0ATKATIDKWCICIKLIFYTSKK 25
    |||||:||||:||||:||||:
DB  782 AMATKAKIDKWDLIKLSFCTAKE 805

RESULT 11
US-10-114-270-30
; Sequence 30, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Bsha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liette, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: NO. US20040030110A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 30
; LENGTH: 1272
; TYPE: PRT
; ORGANISM: Homo sapiens

```


US-10-114-270-30

Query Match 54.9%; Score 79; DB 15; Length 1272;
Best Local Similarity 66.7%; Pred. No. 0.0059;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNKIKLIFYTSKK 25
DB 992 AMATKAKIDKWDLIKLSFCTAKE 1015

RESULT 12

US-10-415-615-3
; Sequence 3, Application US/10415615
; Publication No. US20040101943A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: BAUGHN, Maria R.
; APPLICANT: YUE, Henry
; APPLICANT: LU, Yan
; APPLICANT: DING, Li
; APPLICANT: TANG, Y. Tom
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: RAFALIA, April J.A.
; APPLICANT: LAL, Preeti G.

; TITLE OF INVENTION: NUCLEIC ACID MODIFICATION ENZYMES
; FILE REFERENCE: PI-0280 USN
; CURRENT APPLICATION NUMBER: US/10/415,615
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: PCT/US01/46301
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,458
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/255,107
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 1274
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1258887CD1
US-10-415-615-3

Query Match 54.9%; Score 79; DB 16; Length 1274;
Best Local Similarity 66.7%; Pred. No. 0.0059;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNKIKLIFYTSKK 25
DB 1000 AMATKAKIDKWDLIKLSFCTAKE 1023

RESULT 13

US-10-094-749-1795
; Sequence 1795, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1795
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-1795

Query Match 54.2%; Score 78; DB 14; Length 219;
Best Local Similarity 63.0%; Pred. No. 0.0014;
Matches 17; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNKIKLIFYTSKKEAS 28
DB 67 AMATKAKIDKWDLIKLSFGTAKDTIS 93

RESULT 14

US-09-867-550-1160
; Sequence 1160, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1160
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (3)
; OTHER INFORMATION: Wherein Xaa may be any one of Ala or Asp or Gly or Val
US-09-867-550-1160

Query Match 53.5%; Score 77; DB 9; Length 101;
Best Local Similarity 58.3%; Pred. No. 0.00089;
Matches 14; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNKIKLIFYTSKK 25
DB 15 AEATKAKLDNWDICIKLNFCMAKE 38

RESULT 15

US-10-021-660-94
; Sequence 94, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:

```

; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 94
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-94

```

```

Query Match      52.8%; Score 76; DB 14; Length 647;
Best Local Similarity 66.7%; Pred. NO. 0.0084;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      2 AQATKATIDKNCIKIKIFYTSKK 25
Db      12 AFATKAKIDKWDLIKLSFCTAKE 35

```

```

Search completed: November 11, 2004, 02:43:16
Job time : 33.4349 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 6.86792 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-60

Perfect score: 144

Sequence: 1 IAQATKATIDKWNCKIKLIFYTSKKKAS 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416.

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	54.9	562	2 JU0033	hypothetical L1 pr
2	79	54.9	712	2 S23650	retrovirus-related
3	79	54.9	1275	2 B28096	line-1 protein ORF
4	79	54.9	1280	2 B34087	hypothetical prote
5	73	50.7	1275	2 I38588	reverse transcript
6	73	50.7	1275	2 S85824	reverse transcript
7	72	50.0	1259	4 GNRULL	retrovirus-related
8	69	47.9	92	2 F41925	hypothetical prote
9	56	38.9	275	2 S21348	probable pol polyp
10	56	38.9	513	2 S21976	probable RNA-direc
11	52	36.1	1281	1 GNSULL	retrovirus-related
12	50	34.7	184	2 T38315	very hypothetical
13	50	34.7	806	2 F69899	phage-related pre-
14	49	34.0	488	1 QXASBI	mRNA maturase bil
15	48	33.3	354	2 AE2138	hypothetical prote
16	48	33.3	392	2 S57662	hypothetical prote
17	47.5	33.0	545	2 D30159	hypothetical prote
18	47.5	33.0	2712	2 T30949	hypothetical prote
19	47.5	33.0	3864	2 D37757	protein C4484.1a [
20	47	32.6	206	2 UC2574	hypothetical 22k p
21	46	31.9	1260	4 GNLRLI	retrovirus-related
22	45.5	31.6	94	2 B64550	hypothetical prote
23	44.5	31.6	94	2 F71959	hypothetical prote
24	45.5	31.6	508	1 S62599	UTP-glucose-1-phos
25	45.5	31.6	508	1 JX0277	UTP-glucose-1-phos
26	45.5	31.6	508	2 S35692	UTP-glucose-1-phos
27	45	31.2	129	2 D64444	hypothetical prote
28	45	31.2	485	2 G37324	probable membrane
29	45	31.2	494	2 T05302	vacuolar processin

RESULT 1

JU0033

hypothetical L1 protein (third intron of gene TS) - human

C;Species: Homo sapiens (man)

C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004

C;Accession: JU0033

R;Horie, N.; Nalbantoglu, J.; Kaneda, S.; Ayusawa, D.; Seno, T.; Takeishi, K.

J. Biochem. 106 1-4, 1989

A;Title: Identification and characterization of an L1 family sequence with a very long

A;Reference number: JU0033; MUID:89380111; PMID:2476429

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-562 <HOR>

A;Cross-references: UNIPROT:Q00378

A;Note: this sequence is similar to human teratocarcinoma L1 RNA species and RNA depend

C;Superfamily: pol polyprotein

Query Match 54.9%; Score 79; DB 2; Length 562;
Best Local Similarity 66.7%; Pred. No. 0.00048;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 AQATKATIDKWNCKIKLIFYTSKK 25

Db 489 AMATKAKIDKMDLIKLSFCTAKE 512

RESULT 2

S23650

retrovirus-related hypothetical protein II - human retrotransposon LINE-1

C;Species: Homo sapiens (man)

C;Date: 22-Nov-1993 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004

C;Accession: S23650

R;Hohjoh, H.; Minakami, R.; Sakaki, Y.

Nucleic Acids Res. 18, 4099-4104, 1990

A;Title: Selective cloning and sequence analysis of the human L1 (LINE-1) sequences whi

A;Reference number: S23649; MUID:90332398; PMID:2165587

A;Accession: S23650

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-712 <HOH>

A;Cross-references: UNIPROT:Q15604; UNIPROT:Q9UN80; UNIPROT:Q12881; UNIPROT:Q00363; UNI

PROT:Q14754; UNIPROT:Q9Y5K0; UNIPROT:Q00366; UNIPROT:Q8E30; UNIPROT:Q00375; EMBL:X522;

C;Genetics:

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1990

A;Mobile element: LINE-1

A;Start codon: GTG

C;Superfamily: pol polyprotein

Query Match 54.9%; Score 79; DB 2; Length 712;
Best Local Similarity 66.7%; Pred. No. 0.0006;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 AQATKATIDKWNCKIKLIFYTSKK 25

Db 489 AMATKAKIDKMDLIKLSFCTAKE 512

hypothetical protein 2 - human
C:Species: Homo sapiens (man)
C:date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: F41925
R:Dotzlaw, H.; Alkhalaf, M.; Murphy, L. C.

Mol. Endocrinol. 6, 773-785, 1992
 A;Title: Characterization of estrogen receptor variant mRNAs from human breast cancers.
 A;Reference number: A41925; MUID:92293154; PMID:1603086
 A;Accession: F41925
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-92 <DOT>
 A;Cross-references: UNIPROT:Q14269; GB:M69297; NID:g182218; PIDN:AAA58463.1; PID:g182220
 A;Note: sequence extracted from NCBI backbone (NCBIP:106603)
 C;Superfamily: pol polypeptide

Query Match 47.9%; Score 69; DB 2; Length 92;
 Best Local Similarity 73.7%; Pred. No. 0.0029;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AQAATKATIDKNCIKLKIF 20
 :|||:||||:||||:
 Db 12 AAATKAKIDKDLKIKSP 30

RESULT 9
 S21348
 probable pol polypeptide-related protein 4 - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
 C;Accession: S21348
 R;Schmitz, E.; Moar, E.
 submitted to the EMBL Data Library, June 1990
 A;Reference number: S21345
 A;Accession: S21348
 A;Molecule type: DNA
 A;Residues: 1-275 <SCH>
 A;Cross-references: UNIPROT:Q63306; EMBL:X53581; NID:g56586; PIDN:CAA37647.1; PID:g56590
 C;Superfamily: pol polypeptide
 C;Keywords: polypeptide

Query Match 38.9%; Score 56; DB 2; Length 275;
 Best Local Similarity 45.8%; Pred. No. 0.76;
 Matches 11; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 IAQATKATIDKNCIKLKIFYSK 24
 :|||:||||:||||:
 Db 1 MAYALRSRIDKWDLIKQSFCKAK 24

RESULT 10
 S21976
 probable RNA-directed DNA polymerase (EC 2.7.7.49) (clone MH2C) - rat retrotransposon L1
 N;Alternate names: reverse transcriptase
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 19-Mar-1997 #sequence_revision 17-Oct-1997 #text_change 09-Jul-2004
 C;Accession: S21976
 R;Kahre, O.; Ilves, H.; Speck, M.
 submitted to the EMBL Data Library, August 1991
 A;Reference number: S16783
 A;Accession: S21976
 A;Molecule type: mRNA
 A;Residues: 1-513 <KAH>
 A;Cross-references: UNIPROT:Q63289; EMBL:X61295; NID:g56521; PIDN:CAA43593.1; PID:g56522
 A;Experimental source: clone MH2C
 C;Genetics:
 A;Mobile element: retrotransposon L1
 C;Superfamily: pol polypeptide
 C;Keywords: nucleotidyltransferase; polypeptide; reverse transcriptase

Query Match 38.9%; Score 56; DB 2; Length 513;
 Best Local Similarity 45.8%; Pred. No. 1.4; Mismatches 7; Indels 0; Gaps 0;
 Matches 11; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 IAQATKATIDKNCIKLKIFYSK 24
 :|||:||||:||||:
 Db 442 MAYALRSRIDKWDLIKQSFCKAK 465

RESULT 11

GNMSLL
 retrovirus-related reverse transcriptase homolog - mouse retrotransposon
 N;Alternate names: L1MD repetitive element ORF-2; LINE-1 hypothetical protein; ORF 3900
 C;Species: Mus musculus (house mouse)
 C;Date: 31-Mar-1989 #sequence_revision 08-Jan-1999 #text_change 09-Jul-2004
 C;Accession: B59927; B24906; I49130; A23772; B23430
 R;Loeb, D.D.; Padgett, R.W.; Hardies, S.C.; Shehee, W.R.; Comer, M.B.; Edgell, M.H.; Hu
 Mol. Cell. Biol. 6, 168-182, 1986
 A;Title: The sequence of a large L1MD element reveals a tandemly repeated 5' end and se
 A;Reference number: A93072; MUID:87064284; PMID:3023821
 A;Accession: B59927
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1281 <LOB1>
 A;Cross-references: UNIPROT:P11369; UNIPROT:Q60713; UNIPROT:Q61787; GB:M13002; NID:g200
 A;Note: sequence constructed using the first potential start codon for ORF2

A;Accession: B24906
 A;Molecule type: DNA
 A;Residues: 'NNQESHSTNQKEDSHKNR', 1-1281 <LOE2>
 A;Cross-references: GB:M13002; NID:g200849
 A;Note: sequence shown in Fig. 2
 R;Martin, S.L.; Martin, S.L.
 Gene 153, 261-266, 1995
 A;Title: Characterization of a LINE-1 cDNA that originated from RNA present in ribonuc
 A;Reference number: I49129; MUID:95180729; PMID:7533116
 A;Accession: I49130
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-85, 'L', 87-358, 'K', 360-706, 'F', 708-735, 'A', 737-760, 'W', 762-927, 'D', 929-128
 A;Cross-references: EMBL:U15647; NID:g558906; PIDN:AAA67727.1; PID:g558908
 R;Wotetz, E.; Rogan, P.K.; Manuelidis, L.
 Nucleic Acids Res. 14, 3119-3136, 1986
 A;Title: Conservation in the 5' region of the long interspersed mouse L1 repeat: implic
 A;Reference number: A23772; MUID:86176789; PMID:3008107
 A;Accession: A23772
 A;Molecule type: DNA
 A;Residues: 'NNQESHSTNQKEDSHKNR', 1-245, 'K', 247-423, 'SYTQONWKTWTKWTN', 439, 'WTDTRYQS', <W

Query Match 36.1%; Score 52; DB 1; Length 1281;
 Best Local Similarity 41.7%; Pred. No. 13;
 Matches 10; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy 1 IAQATKATIDKNCIKLKIFYSK 24
 :|||:||||:||||:
 Db 1007 MACAVRSRIDKWDLMKLSQFCKAK 1030

RESULT 12

T38315
 very hypothetical protein SPAC23H4.13c - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C;Accession: T38315
 R;Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1997
 A;Reference number: Z21733
 A;Accession: T38315
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-184 <BRO>
 A;Cross-references: UNIPROT:O13954; EMBL:Z98977; PIDN:CAB11667.1; GSPDB:GN00066; SPDB:S
 A;Experimental source: strain 972h-; cosmid c23H4
 C;Genetics:
 A;Gene: SPDB:SPAC23H4.13c
 A;Map position: 1
 A;Introns: 57/3
 C;Superfamily: Schizosaccharomyces pombe very hypothetical protein SPAC23H4.13c

Query Match 34.7%; Score 50; DB 2; Length 184;
 Best Local Similarity 66.7%; Pred. No. 4.2; Mismatches 3; Indels 1; Gaps 0;
 Matches 8; Conservative 3; Mismatches 3; Indels 1; Gaps 0;

QY 8 TTKWNCIKIKI 19
 |||:|||||:
 Db 104 TTKWNCIKIKI 115

RESULT 13
 F69899
 phage-related pre-neck appendage protein homolog yobO - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: F69899
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Erington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallex
 A:Authors: Schleich, S.; Schroefer, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 ieck, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Tauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroefer, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumstein, B.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:198044033; PMID:9384377
 A:Accession: F69899
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-806 <UNP>
 A:Cross-references: UNIPROT:Q34433; GB:Z99114; GB:AL009126; NID:G2634230; PIDN:CAB13795.
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yobO
 C:Superfamily: phage PZA gene 12 protein

Query Match 34.7%; Score 50; DB 2; Length 806;
 Best Local Similarity 47.6%; Pred. No. 17;
 Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 TTKATDKWNCIKIKIFYTSKK 25
 |||:|||||:
 Db 776 TTKATDKWNCIKIKIFYTSKK 796

RESULT 14
 QXASBI
 mRNA maturase bil - Emericella nidulans mitochondrion
 N:Alternate names: cOXA intron protein
 C:Species: Mitochondrion Emericella nidulans, Aspergillus nidulans
 C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
 C:Accession: A04513; A37516
 R:Waring, R.B.; Davies, R.W.; Lee, S.; Grisi, E.; Berks, M.M.; Sczzocchio, C.
 Cell 27, 4-11, 1981
 A:Title: The mosaic organization of the apocytochrome b gene of Aspergillus nidulans rev
 A:Reference number: A00157; MUID:82115341; PMID:7034966
 A:Accession: A04513
 A:Molecule type: DNA
 A:Residues: 1-237 <WAL>
 A:Cross-references: UNIPROT:P03880; GB:J01388; GB:V00651; GB:V00652; NID:G336
 A:Experimental source: imperfect stage
 A:Note: This ORF is not annotated in GenBank entry EMENTCYB2, release 106.0
 R:Waring, R.B.; Davies, R.W.; Sczzocchio, C.; Brown, T.A.
 Proc. Natl. Acad. Sci. U.S.A. 79, 6332-6336, 1982
 A:Title: Internal structure of a mitochondrial intron of Aspergillus nidulans.
 A:Reference number: A37516; MUID:83065170; PMID:6755468
 A:Accession: A37516
 A:Molecule type: DNA

A:Residues: 161-488 <WA2>
 A:Cross-references: GB:J01387; NID:G336899
 A:Experimental source: imperfect stage
 A:Note: This ORF is not annotated in GenBank entry EMENTCOBA, release 106.0
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC3
 C:Superfamily: Emericella mRNA maturase bil; cytochrome b homology; cytochrome b6 homol
 C:Keywords: mitochondrion
 F:1-168/Region: cytochrome b exon encoded
 F:10-168/Domain: cytochrome b homology (fragment) <CBH>
 F:10-168/Domain: cytochrome b6 homology (fragment) <CB6>
 F:169-488/Region: cytochrome b intron encoded

Query Match 34.0%; Score 49; DB 1; Length 488;
 Best Local Similarity 60.0%; Pred. No. 15;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 TTKATDKWNCIKIKI 19
 |||:|||||:
 Db 423 TTKVLDKTNCSKLKV 437

RESULT 15
 AE2138
 hypothetical protein alr2660 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AE2138
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AE2138
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-354 <KUR>
 A:Cross-references: UNIPROT:Q8YQ3; GB:BA000019; PIDN:BA074359.1; PID:gl7131753; GSPDB:
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr2660
 C:Superfamily: uncharacterized protein with thioresdoxin-like domain

Query Match 33.3%; Score 48; DB 2; Length 354;
 Best Local Similarity 40.0%; Pred. No. 16;
 Matches 10; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 6 KATIDKWN---CIKLKIFYSKKE 26
 |||:|||||:
 Db 289 KLLVERNEQATWVKLAISYTSKDD 313

Search completed: November 10, 2004, 14:52:27
 Job time : 7.86792 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 37.2981 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-60
 Perfect score: 144
 Sequence: 1 IAOATKATIDKWNICIKLIFYTSKKEAS 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters:	1825181
--	---------

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45

```

```
Database :      UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	144	100.0	28	2	Q96JA8	Q96ja8 homo sapien
2	79	54.9	254	2	Q62U70	Q62u70 homo sapien
3	79	54.9	254	2	BAC86356	BAC86356 homo sapi
4	79	54.9	473	2	Q8NB08	Q8nb08 homo sapien
5	79	54.9	712	2	Q14754	Q14754 homo sapien
6	79	54.9	1275	2	Q00360	Q00360 homo sapien
7	79	54.9	1275	2	Q00362	Q00362 homo sapien
8	79	54.9	1275	2	Q00372	Q00372 homo sapien
9	79	54.9	1275	2	Q00375	Q00375 homo sapien
10	79	54.9	1275	2	Q00378	Q00378 homo sapien
11	75	52.1	164	2	Q62VR1	Q62vr1 homo sapien
12	75	52.1	164	2	BAC85800	BAC85800 homo sapi
13	75	52.1	1275	2	Q8TE30	Q8te30 homo sapien
14	73	50.7	1275	2	Q00549	Q00549 homo sapien
15	73	50.7	1192	2	Q7K241	Q7k241 homo sapien
16	73	50.7	1275	2	Q00363	Q00363 homo sapien
17	73	50.7	1275	2	Q00366	Q00366 homo sapien
18	73	50.7	1275	2	Q00368	Q00368 homo sapien
19	73	50.7	1275	2	Q00370	Q00370 homo sapien
20	72	50.0	122	2	Q8WY74	Q8wy74 homo sapien
21	72	50.0	126	2	Q8N1K2	Q8n1k2 homo sapien
22	72	50.0	1259	1	LI_N1_HUMAN	P08547 homo sapien
23	71	49.3	206	2	Q8M2I2	Q8m2i2 homo sapien
24	71	49.3	644	2	Q8N8P1	Q8n8p1 homo sapien
25	67	46.5	1210	2	Q8EQJ8	Q8eqj8 oceanobacil
26	65	45.1	1102	2	Q6TUD6	Q6tude rattus norv
27	65	45.1	1102	2	AAQ91064	AAQ91064 rattus no
28	58	40.3	1446	2	Q7TP07	Q7tp07 rattus norv
29	56	38.9	275	2	Q63306	Q63306 rattus norv
30	56	38.9	370	2	Q6TUG6	Q6tug6 rattus norv
31	56	38.9	370	2	AAQ91034	AAQ91034 rattus no

32	56	38.9	431	2	Q6TXG5	Q6txg5 rattus norv
33	56	38.9	431	2	AQ936247	Aq936247 rattus no
34	56	38.9	513	2	Q63289	Q63289 rattus norv
35	56	38.9	621	2	Q6Q153	Q6q153 rattus norv
36	56	38.9	621	2	AAS66246	Aas66246 rattus no
37	56	38.9	696	2	Q6Q137	Q6q137 rattus norv
38	56	38.9	696	2	AAS66262	Aas66262 rattus no
39	56	38.9	753	2	Q6TXG3	Q6txg3 rattus norv
40	56	38.9	753	2	AQ936249	Aq936249 rattus no
41	56	38.9	948	2	Q7TFL1	Q7tfl1 rattus norv
42	56	38.9	1015	2	Q7TNM9	Q7tnm9 rattus norv
43	56	38.9	1073	2	Q6Q157	Q6q157 rattus norv
44	56	38.9	1073	2	AAS66242	Aas66242 rattus no
45	56	38.9	1104	2	Q6TUH9	Q6tuh9 rattus norv

ALIGNMENTS

RESULT 1

Q96JA8		PRELIMINARY;	PRT;	28 AA.
ID	Q96JA8			
AC	Q96JA8			
DT	01-DEC-2001	(TEMBRel. 19, Created)		
DT	01-DEC-2001	(TEMBRel. 19, Last sequence update)		
DT	01-DEC-2001	(TEMBRel. 19, Last annotation update)		
DE	Bcl-Xl-binding protein c21 (Fragment).			
DE	Bcl-Xl-binding protein c21 (Fragment).			
OS	Homo sapiens (Human).			
CS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=kidney;			
RC	MEDLINE=21293069; PubMed=11283018;			
RA	Hammond P.W., Alpin J., Rise C.E., Wright M., Kreider B.L.;			
RT	"In Vitro Selection and Characterization of Bcl-Xl-Binding Proteins			
RT	from a Mix of Tissue-specific mRNA Display Libraries.";			
RL	J. Biol. Chem. 276:20898-20906(2001).			
DR	EWBL; AF357525; AAK60629.1; -.			
FT	NON TER	1		
FT	NON TER	1		
FT	NON TER	28		
SO	SEQUENCE	28 AA;	3201 MW;	A7DCE42AA280631E CRC64;

June 15

RESULT 2

Q6ZU70 PRELIMINARY; PRT; 254 AA.

ID Q6ZU70

AC Q6ZU70;

DT 05-JUL-2004 (T-EMBLrel. 27, Created)

DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)

DE Hypothetical protein FLJ43960.

DE Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

NCBI_TaxId=9606;

FN [1]

RN SEQUENCE FROM N.A.

RP

RC TISSUE=Testis;

RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,

RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,

RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,

RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,

RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,

```

RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.,
RA Submitted (JUL-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AK125948; BAC86356.1; -.
SQ SEQUENCE 254 AA; 30123 MW; AFE89F7AB72FA49D CRC64;

Query Match          54.9%; Score 79; DB 2; Length 254;
Best Local Similarity 66.7%; Pred. No. 0.00075;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQATKATIDKWNCKIKLIFYTSKK 25
Db 12 AMATKAKIDKWLKIKSFCSTAKE 35

RESULT 3
ID BAC86356 PRELIMINARY; PRT; 254 AA.
AC BAC86356;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE CDNA FLJ43960 fis, clone T5ST14016551.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Testis;
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato K., Nishikawa T., Ishii S., Yamamoto Y., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RA "NEDO human cDNA sequencing project.";
RT Submitted (JUL-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AK125948; BAC86356.1; -.
SQ SEQUENCE 254 AA; 30123 MW; AFE89F7AB72FA49D CRC64;

Query Match          54.9%; Score 79; DB 2; Length 254;
Best Local Similarity 66.7%; Pred. No. 0.00075;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQATKATIDKWNCKIKLIFYTSKK 25
Db 12 AMATKAKIDKWLKIKSFCSTAKE 35

RESULT 4
ID Q8NB08 PRELIMINARY; PRT; 473 AA.
AC Q8NB08;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ34421.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=14702039;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,

```

```

RA Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa B., Omura Y.,
RA Abe K., Kanihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotsuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Koriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senba T.,
RA Katsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu I.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RT Nat. Genet. 36:40-45(2004).
RL EMBL; AK091740; BAC03736.1; -.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; RVT; 1.
DR KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 473 AA; 55258 MW; 5DD23CF8849FABDA CRC64;

Query Match          54.9%; Score 79; DB 2; Length 473;
Best Local Similarity 66.7%; Pred. No. 0.0014;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQATKATIDKWNCKIKLIFYTSKK 25
Db 400 AMATKAKIDKWLKIKSFCSTAKE 423

RESULT 5
ID Q14754 PRELIMINARY; PRT; 712 AA.
AC Q14754;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF11..
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Japanese; TISSUE=Placenta;
RX MEDLINE=90332398; PubMed=21655587;
RA Honjoh H., Minakami R., Sakaki Y.;
RT "Selective cloning of the Human LI (line-1) sequence which transposed
RT in a relatively recent past.";
RL Nucleic Acids Res. 18:4099-4104(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Japanese; TISSUE=Placenta;
RX MEDLINE=92319645; PubMed=1320255;
RA Minakami R., Kurase K., Etoh K., Furuhata Y., Hattori M., Sakaki Y.;
RT "Identification of an internal cis-element essential for the human LI
RT transcription and a nuclear factor(s) binding to the element.";
RL Nucleic Acids Res. 20:3139-3145(1992).

```



```

RN RP SEQUENCE FROM N.A.
RN RP MEDLINE=97285120; PubMed=9140393;
RA Sessaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human L1 elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL, U93564; AAC51263.1; -.
DR FIR, B28096; S28096.
DR FIR, S23650; S23650.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVT; 1.
DR RNA-directed DNA polymerase; Transferase.
DR KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1275 AA; 149201 MW; 23D516D6E4358F28 CRC64;

Query Match 54.9%; Score 79; DB 2; Length 1275;
Best Local Similarity 66.7%; Pred. No. 0.0038; 5; Indels 0; Gaps 0;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNCKIKIFYTSKK 25
DB 1001 AWATKAKIDKWDLIKLSFCTAKE 1024
|||||:|||||:|||||:|:|:|

RESULT 8
O00372 PRELIMINARY; PRT; 1275 AA.
AC O00372;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative p150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI TaxID=9606;
RN [1]_TaxID=9606;
RN RP SEQUENCE FROM N.A.
RN RP MEDLINE=97285120; PubMed=9140393;
RA Sessaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human L1 elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL, U93570; AAC51273.1; -.
DR FIR, B28096; S28096.
DR FIR, S23650; S23650.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVT; 1.
DR RNA-directed DNA polymerase; Transferase.
DR KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1275 AA; 149110 MW; 4711B3BC22F7674E CRC64;

Query Match 54.9%; Score 79; DB 2; Length 1275;
Best Local Similarity 66.7%; Pred. No. 0.0038; 5; Indels 0; Gaps 0;
Matches 16; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 AQAATKATIDKWCNCKIKIFYTSKK 25
DB 1001 AWATKAKIDKWDLIKLSFCTAKE 1024
|||||:|||||:|||||:|:|:|

RESULT 9
O00375 PRELIMINARY; PRT; 1275 AA.
ID O00375

```

Page 4

```

Qy  4  ATKATIDKNCIKLKIFYTSKK 25
      ||||| ||||| ||||| |||||
Db  9  ATXAKIDKNDLIKLSFCTAKE 30

RESULT 13
Q8TE30
ID  Q8TE30      PRELIMINARY;      PRT; 1275 AA.
AC  Q8TE30;
DT  01-JUN-2002 (TrEMBLrel. 21, Created)
DT  01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Hypothetical protein.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21668188; PubMed=11810275;
RA  Banjes S.M., Morris C.M.
RT  "A full-length and potentially active LINE element is integrated
RT  polymorphically within the IGL locus in a genomically unstable region
RT  of chromosome 22."
RL  Hum. Genet. 109:628-637(2001).
DR  EMBL; AF421375; AAL50637.1; -.
DR  PIR; B28096; B28096.
DR  PIR; S23650; S23650.
DR  GO; GO:0003723; F:RNA binding; IEA.
DR  GO; GO:000364; F:RNA-directed DNA polymerase activity; IEA.
DR  GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR  InterPro; IPR0051135; Exo_endo_phos.
DR  InterPro; IPR000477; RVTse.
DR  Pfam; PF03372; Exo_endo_phos; 1.
DR  Pfam; PF00078; RVT; 1.
KW  Hypothetical protein; RNA-directed DNA polymerase; Transferase.
SQ  SEQUENCE 1275 AA; 149009 MW; B327D9D50A581764 CRC64;

Query Match      52.1%; Score 75; DB 2; Length 1275;
Best Local Similarity 62.5%; Pred. No. 0.016;
Matches 15; Conservative 3; Mismatches 6; Indels 0; Gaps

Qy  2  AQATKATIDKNCIKLKIFYTSKK 25
      ||||| ||||| ||||| |||||
Db  1001  AMATKVIDKNDLIKLSFCTAKE 1024

RESULT 14
Q00549
ID  Q00549      PRELIMINARY;      PRT; 573 AA.
AC  Q00549;
DT  01-JUL-1997 (TrEMBLrel. 04, Created)
DT  01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  ORF2-like protein (Fragment).
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  TISSUE=Small-cell lung cancer;
RX  MEDLINE=9918768; PubMed=9918768;
RA  Kuo K.W., Shou H.M., Huang Y.S., Leung W.C.;
RT  "Expression of transposon LINE-1 is relatively human-specific and
RT  function of the transcripts may be proliferation-essential."
RL  Biochem. Biophys. Res. Commun. 253:566-570(1998).
DR  EMBL; AF003535; AAD04635.1; -.
DR  GO; GO:0003723; F:RNA binding; IEA.
DR  GO; GO:000364; F:RNA-directed DNA polymerase activity; IEA.
DR  GO; GO:0016740; F:transferase activity; IEA.
DR  GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR  InterPro; IPR000477; RVTse.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 8.05283 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-61

Perfect score: 113

Sequence: 1 VDVDPFIVWLEAVSDLRAL 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	41.6	557	4	US-09-138-277C-3
2	46	40.7	126	4	US-09-107-532A-4794
3	46	40.7	304	1	US-08-569-806-6
4	45.5	40.3	354	4	US-08-252-991A-27916
5	45	39.8	365	4	US-08-328-352-6107
6	45	39.8	405	4	US-09-270-767-44068
7	44	38.9	299	4	US-08-252-991A-18741
8	44	38.9	336	4	US-09-489-039A-9200
9	44	38.9	351	4	US-08-538-092-1255
10	44	38.9	1003	4	US-08-489-039A-12357
11	43.5	38.5	691	3	US-08-134-001C-4675
12	43	38.1	171	4	US-09-489-039A-9952
13	43	38.1	375	4	US-09-540-236-2085
14	42.5	37.6	164	4	US-09-107-532A-6548
15	42	37.2	37	4	US-08-716-129-100
16	42	37.2	145	4	US-08-716-129-96
17	42	37.2	180	4	US-08-543-681A-7612
18	42	37.2	205	4	US-08-489-039A-11450
19	42	37.2	167	4	US-08-489-039A-12889
20	42	37.2	282	4	US-08-270-767-41601
21	42	37.2	481	3	US-08-537-357-7
22	42	37.2	487	3	US-08-537-357-15
23	42	37.2	943	3	US-08-327-885-5
24	42	37.2	943	4	US-08-969-362-5
25	42	37.2	946	4	US-08-492-709A-326
26	41.5	36.7	404	4	US-08-489-847-141
27	41.5	36.7	489	4	US-08-489-847-289

28	41	36.3	71	3	US-09-073-297-10	Sequence 10, Appl
29	41	36.3	117	4	US-09-270-767-39094	Sequence 39094, A
30	41	36.3	117	4	US-09-270-767-54311	Sequence 54311, A
31	41	36.3	266	4	US-09-252-991A-31767	Sequence 31767, A
32	41	36.3	330	4	US-09-489-039A-13171	Sequence 13171, A
33	41	36.3	361	4	US-09-540-236-2502	Sequence 2502, Ap
34	41	36.3	438	4	US-09-489-039A-13834	Sequence 13834, A
35	41	36.3	476	3	US-09-188-579-114	Sequence 114, App
36	41	36.3	476	3	US-09-315-444-114	Sequence 114, App
37	41	36.3	476	3	US-09-134-218-6	Sequence 6, Appl
38	41	36.3	476	4	US-09-721-362-114	Sequence 114, Appl
39	41	36.3	762	2	US-08-907-166-10	Sequence 10, Appl
40	41	36.3	762	4	US-09-391-340-10	Sequence 10, Appl
41	41	36.3	983	4	US-09-252-991A-32868	Sequence 32868, A
42	41	36.3	2987	2	US-08-970-269A-29	Sequence 29, Appl
43	41	36.3	2987	3	US-09-407-562-29	Sequence 29, Appl
44	41	36.3	3959	2	US-08-970-269A-30	Sequence 30, Appl
45	41	36.3	3959	3	US-09-407-562-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1

US-09-138-277C-3
; Sequence 3, Application US/09138277C
; Patent No. 6436403
; GENERAL INFORMATION:
; APPLICANT: NAKATA, MOTOMI
; APPLICANT: NAKANO, HIROYASU
; APPLICANT: YAGITA, HIDEO
; APPLICANT: OKUMURA, KO
; TITLE OF INVENTION: TRAP FAMILY MOLECULES, POLYNUCLEOTIDES ENCODING THEM,
; TITLE OF INVENTION: AND ANTIBODIES AGAINST THEM
; FILE REFERENCE: 007898-0255515
; CURRENT APPLICATION NUMBER: US/09/138,277C
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: PCT/JP97/00512
; PRIOR FILING DATE: 1997-02-24
; PRIOR APPLICATION NUMBER: JP 34674/1996
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-138-277C-3

Query Match 41.6%; Score 47; DB 4; Length 557;
Best Local Similarity 64.3%; Pred. No. 38;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VMLEAVSDLRAL 22
Db 258 VLEEQISDLHKS 271

RESULT 2

US-09-107-532A-4794
; Sequence 4794, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA

```

; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4794:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...126
; SEQUENCE DESCRIPTION: SEQ ID NO: 4794:
US-09-107-532A-4794

Query Match          40.7%; Score 46; DB 4; Length 126;
Best Local Similarity 58.3%; Pred. No. 10;
Matches              7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      9 VWLEAVSDLHR 20
DB      110 IWIEEALSELR 121
        :|::|||:|:|
        :|::|||:|:|

RESULT 3
US-08-569-806-6
; Sequence 6, Application US/08569806
; Patent No. 5721126
; GENERAL INFORMATION:
; APPLICANT: XU, SHUANG-YONG
; APPLICANT: XIAO, JIAN-PING
; TITLE OF INVENTION: METHOD FOR CLONING AND
; TITLE OF INVENTION: PRODUCING THE ScaI RESTRICTION ENDONUCLEASE IN
; TITLE OF INVENTION: E. COLI
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND
; ADDRESSER: BIOLABS, INC.
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/569,806
; FILING DATE:
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 304 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-569-806-6

Query Match          40.7%; Score 46; DB 1; Length 304;
Best Local Similarity 38.1%; Pred. No. 28;
Matches              8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY      1 VVDVPDFIVWLEAEVSDLHRA 21
DB      88 VLISERYISWCETWMSRVHRA 108
        |.:.:|.|.:|.:|
        |.:.:|.|.:|.:|

RESULT 4
US-09-252-991A-27916
; Sequence 27916, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27916
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-27916

Query Match          40.3%; Score 45.5; DB 4; Length 354;
Best Local Similarity 66.7%; Pred. No. 39;
Matches             10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      4 VPDPFIWLEAEVSDL 18
DB      41 VPSF-AWAEDAVSDL 54
        |||:|||:|
        |||:|||:|

RESULT 5
US-09-328-352-6107
; Sequence 6107, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6107
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-6107

```

Query Match 39.8%; Score 45; DB 4; Length 365;
 Best Local Similarity 61.5%; Pred. No. 48;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 DFLVLEEAASDL 18
 ||:|||||:
 Db 199 DFLVLEENMDGL 211

RESULT 6
 US-09-270-767-44068
 ; Sequence 44068, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 44068
 ; LENGTH: 405
 ; TYPE: PRT
 ; ORGANISM: *Drosophila melanogaster*
 US-09-270-767-44068

Query Match 39.8%; Score 45; DB 4; Length 405;
 Best Local Similarity 57.1%; Pred. No. 54;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 IWLLEEAASDLHRA 21
 ||:|||||:
 Db 383 IWLIAECVAQKHRA 396

RESULT 7
 US-09-252-991A-18741
 ; Sequence 18741, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 18741
 ; LENGTH: 299
 ; TYPE: PRT
 ; ORGANISM: *Pseudomonas aeruginosa*
 US-09-252-991A-18741

Query Match 38.9%; Score 44; DB 4; Length 299;
 Best Local Similarity 75.0%; Pred. No. 55;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 LEEAASDLHRL 22
 :|||||:
 Db 76 VEEAVAGLHRL 87

RESULT 8
 US-09-489-039A-9200
 ; Sequence 9200, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 9200
 ; LENGTH: 336
 ; TYPE: PRT
 ; ORGANISM: *Klebsiella pneumoniae*
 US-09-489-039A-9200

Query Match 38.9%; Score 44; DB 4; Length 336;
 Best Local Similarity 47.1%; Pred. No. 63;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVDFIVWLEEAASD 17
 ||:|||||:
 Db 30 VMKVSDDLTLAEAILD 46

RESULT 9
 US-09-538-092-1256
 ; Sequence 1256, Application US/09538092
 ; Patent No. 6753314
 ; GENERAL INFORMATION:
 ; APPLICANT: Giot, Loic
 ; APPLICANT: Mansfield, Traci A.
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
 ; FILE REFERENCE: 15966-542
 ; CURRENT APPLICATION NUMBER: US/09/538,092
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/127,352
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178,965
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 1387
 ; SOFTWARE: CuratSeqformatter Version 0.9
 ; SEQ ID NO 1256
 ; LENGTH: 351
 ; TYPE: PRT
 ; ORGANISM: *Homo sapiens*
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (0)...(0)
 ; OTHER INFORMATION: Polypeptide Accession Number Q02547
 US-09-538-092-1256

Query Match 38.9%; Score 44; DB 4; Length 351;
 Best Local Similarity 50.0%; Pred. No. 66;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 IWLLEEAASDLHRA 21
 ||:|||||:
 Db 329 IWLIAECIAQRHRA 342

RESULT 10
 US-09-489-039A-12357
 ; Sequence 12357, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342

```
; SEQ ID NO 12357
; LENGTH: 1003
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12357

Query Match      38.9%; Score 44; DB 4; Length 1003;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 VDVDPFVWLEEA VSDLHRL 22
Db 313 VVDEAHLVWSEAFSREYQAI 334

RESULT 11
US-09-134-001C-4675
; Sequence 4675, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4675
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4675

Query Match      38.5%; Score 43.5; DB 3; Length 691;
Best Local Similarity 45.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 4 VPDFIV-WLEEA VSDLHRL 22
Db 357 IPDFVLEQLFHEHNDLHKL 376

RESULT 12
US-09-489-039A-9952
; Sequence 9952, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9952
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9952

Query Match      38.1%; Score 43; DB 4; Length 171;
Best Local Similarity 44.4%; Pred. No. 43;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 DVPDFVWLEEA VSDLHR 20
Db 67 DLPDSIVPLYEGIKDMYQ 104

RESULT 13
US-09-540-236-2085
; Sequence 2085, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Bretton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2085
; LENGTH: 375
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2085

Query Match      38.1%; Score 43; DB 4; Length 375;
Best Local Similarity 42.9%; Pred. No. 1e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 DRIWLEEA VSDLH 19
Db 202 DFLWLEHADQIN 215

RESULT 14
US-09-107-532A-6548
; Sequence 6548, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6548:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
```



```

; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...164
; SEQUENCE DESCRIPTION: SEQ ID NO: 6548:
US-09-107-532A-6548

```

```

Query Match      37.6%; Score 42.5; DB 4; Length 164;
Best Local Similarity 27.6%; Pred. No. 49;
Matches 8; Conservative 8; Mismatches 6; Indels 7; Gaps 1;

```

```

Qy 1 VVDVPDFI-----VMLEEAVSDIHLRAL 22
Db 78 VVEIPAVVGATGVEPITLRKFISDFHKGL 106

```

```

RESULT 15
US-09-716-129-100
; Sequence 100, Application US/09716129
; Patent No. 6632920
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2025P1
; CURRENT APPLICATION NUMBER: US/09/716,129
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/076,053
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,057
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,052
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,054
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,051
; PRIOR FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-716-129-100

```

```

Query Match      37.2%; Score 42; DB 4; Length 37;
Best Local Similarity 36.8%; Pred. No. 11;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 4 VPDFIVMLEEAVSDIHLRAL 22
Db 9 LPDYLAIDEALALHVTI 27

```

Search completed: November 10, 2004, 14:55:43
Job time : 8.10283 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 25.4453 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-61

Perfect score: 113

Sequence: 1 VVDVPDFVWLEAVSDLHRL 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	22	14	US-10-092-750-61
2	51	54.0	22	14	US-10-092-750-227
3	50	44.2	351	14	US-10-369-493-11619
4	50	44.2	355	14	US-10-369-493-14336
5	50	44.2	355	14	US-10-369-493-14603
6	50	44.2	355	14	US-10-369-493-14999
7	49	43.4	790	16	US-10-437-963-156459
8	48	42.5	611	15	US-10-382-122A-69638
9	47	41.6	78	16	US-10-437-963-124299
10	47	41.6	451	14	US-10-004-378A-2
11	47	41.6	538	15	US-10-042-865-163
12	47	41.6	557	14	US-10-004-378A-34
13	47	41.6	557	15	US-10-042-865-162

14	47	41.6	568	15	US-10-262-445-132	Sequence 132, Appl
15	47	41.6	568	15	US-10-042-865-38	Sequence 38, Appl
16	47	41.6	1022	14	US-10-156-761-9292	Sequence 9292, Ap
17	46	40.7	848	9	US-09-882-986-2	Sequence 2, Appl
18	46	40.7	848	15	US-10-469-013-18	Sequence 18, Appl
19	46	40.7	1295	16	US-10-437-963-144134	Sequence 144134,
20	46	40.7	3562	14	US-10-341-434-109	Sequence 109, App
21	45.5	40.3	327	17	US-10-435-115-199784	Sequence 199784,
22	45.5	40.3	541	16	US-10-437-963-151993	Sequence 151993,
23	45	39.8	240	15	US-10-282-122A-47961	Sequence 47961, A
24	45	39.8	266	14	US-10-369-493-8199	Sequence 8199, Ap
25	45	39.8	402	14	US-10-369-493-18774	Sequence 18774, A
26	45	39.8	1004	14	US-10-156-761-14806	Sequence 14806, A
27	44.5	39.4	317	14	US-10-369-493-1258	Sequence 1258, Ap
28	44.5	39.4	317	14	US-10-369-493-20325	Sequence 20325, A
29	44.5	39.4	317	14	US-10-369-493-21652	Sequence 21652, A
30	44.5	39.4	543	14	US-10-369-493-17936	Sequence 17936, A
31	44	38.9	118	15	US-10-424-599-175022	Sequence 175022,
32	44	38.9	129	16	US-10-437-963-165340	Sequence 165340,
33	44	38.9	175	15	US-10-424-599-253697	Sequence 253697,
34	44	38.9	232	15	US-10-276-774-2499	Sequence 2499, Ap
35	44	38.9	234	17	US-10-489-372-23	Sequence 23, Appl
36	44	38.9	345	14	US-10-365-646-4	Sequence 4, Appl
37	44	38.9	345	14	US-10-365-646-5	Sequence 5, Appl
38	44	38.9	351	10	US-09-997-003-33	Sequence 33, Appl
39	44	38.9	400	14	US-10-369-493-20208	Sequence 20208, A
40	44	38.9	461	14	US-10-176-306-23	Sequence 23, Appl
41	44	38.9	461	14	US-10-094-749-2145	Sequence 2145, Ap
42	44	38.9	461	14	US-10-285-027-305	Sequence 205, App
43	44	38.9	461	15	US-10-188-832-97	Sequence 97, Appl
44	44	38.9	906	16	US-10-437-963-105322	Sequence 105322,
45	44	38.9	943	17	US-10-425-115-278341	Sequence 278341,

ALIGNMENTS

RESULT 1
US-10-092-750-61
; Sequence 61, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-61

Query Match 100.0%; Score 113; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VVDVPDFVWLEAVSDLHRL 22
DB 1 VVDVPDFVWLEAVSDLHRL 22
RESULT 2
US-10-092-750-227
; Sequence 227, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:

```

; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 227
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-227

Query Match          54.0%; Score 61; DB 14; Length 22;
Best Local Similarity 75.0%; Pred. No. 0.048;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 DVDPFIVLLEAVSDLHRAL 22
DB 3 DVVGFIDSEGVSDLHRAL 22

RESULT 3
US-10-369-493-11619
; Sequence 11619, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11619
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-11619

Query Match          44.2%; Score 50; DB 14; Length 351;
Best Local Similarity 44.4%; Pred. No. 40;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVPDFIVLLEAVSDL 18
DB 190 LIDKPDFFFWLEKNWDDI 207

RESULT 4
US-10-369-493-14336
; Sequence 14336, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14999
; LENGTH: 355
; TYPE: PRT

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14336
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14336

Query Match          44.2%; Score 50; DB 14; Length 355;
Best Local Similarity 44.4%; Pred. No. 41;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVPDFIVLLEAVSDL 18
DB 188 LIDKPDFFFWLEKNWDDI 205

RESULT 5
US-10-369-493-14603
; Sequence 14603, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14603
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14603

Query Match          44.2%; Score 50; DB 14; Length 355;
Best Local Similarity 44.4%; Pred. No. 41;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVPDFIVLLEAVSDL 18
DB 188 LIDKPDFFFWLEKNWDDI 205

RESULT 6
US-10-369-493-14999
; Sequence 14999, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14999
; LENGTH: 355
; TYPE: PRT

```

```
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14999

Query Match      44.2%; Score 50; DB 14; Length 355;
Best Local Similarity 44.4%; Pred. No. 41;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY      1 VVDVDFIVWLEEA VSDL 18
        :|||||:
Db      188 LIDXPDPFFWLEKWNDDI 205

RESULT 7
US-10-437-963-156459
; Sequence 156459, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 156459
; LENGTH: 790
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56124C.1.pep
US-10-437-963-156459

Query Match      43.4%; Score 49; DB 16; Length 790;
Best Local Similarity 53.3%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      3 DVPDFIVWLEEA VSD 17
        :|||||:
Db      627 DQPDFISWVSECLD 641

RESULT 8
US-10-282-122A-69638
; Sequence 69638, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69638
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69638

Query Match      42.5%; Score 48; DB 15; Length 611;
Best Local Similarity 52.6%; Pred. No. 1.5e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY      3 DVPDFIVWLEEA VSDLHRA 21
        :|||||:
Db      139 DTFDLAVALKSAVKELHGA 157

RESULT 9
US-10-437-963-124299
; Sequence 124299, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 124299
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27051C.1.pap
US-10-437-963-124299

Query Match      41.6%; Score 47; DB 16; Length 78;
Best Local Similarity 69.2%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      9 WLEEA VSDLHRA 21
        :|||||:
Db      41 LWLEEA PSLHRA 53

RESULT 10
US-10-004-378A-2
; Sequence 2, Application US/10004378A
```

```

; Publication No. US20030228301A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Furtak, Kazaryna
; APPLICANT: Perna, Amanda
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corinne A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A
; TITLE OF INVENTION: No. US20030228301A1el Human Proteins, Polynucleotides Encoding TH
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004,378A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-378A-2

Query Match 41.6%; Score 47; DB 14; Length 451;
Best Local Similarity 64.3%; Pred. No. 1.5e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 9 VMLEAVSDLHRL 22
Db 152 VQLEQISDLHKS 165

RESULT 11
US-10-042-865-163

```

```

; Sequence 163, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zernhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corinne A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Schlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 163
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-865-163

Query Match 41.6%; Score 47; DB 15; Length 538;
Best Local Similarity 64.3%; Pred. No. 1.8e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 9 VMLEAVSDLHRL 22
Db 239 VQLEQISDLHKS 252

RESULT 12
US-10-004-378A-34
; Sequence 34, Application US/10004378A
; Publication No. US20030228301A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Furtak, Kazaryna
; APPLICANT: Perna, Amanda
; APPLICANT: Patturajan, Meera

```

```

; APPLICANT: Shimkets, Richard A
; APPLICANT: Guo, Xiaojia Sashia
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corrine A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A
; TITLE OF INVENTION: No. US2003028301A1el Human Proteins, Polynucleotides Encoding Them
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004,378A
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-378A-34

Query Match 41.6%; Score 47; DB 14; Length 557;
Best Local Similarity 64.3%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VMLEEAVSDLHRL 22
DB 258 VQLEEQISDLHKS 271

RESULT 13
US-10-042-865-162
; Sequence 162, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zethusen, Bryan D

```

```

; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Schlomit R
; APPLICANT: Rotherberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Willet, Isabelle
; APPLICANT: Feyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 162
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-865-162

Query Match 41.6%; Score 47; DB 15; Length 557;
Best Local Similarity 64.3%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VMLEEAVSDLHRL 22
DB 258 VQLEEQISDLHKS 271

RESULT 14
US-10-262-445-132
; Sequence 132, Application US/10262445
; Publication No. US20040014058A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John
; APPLICANT: Burgess, Catherine
; APPLICANT: Catterton, Elina
; APPLICANT: Chant, John
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Edinger, Schlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Giot, Loic
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia

```

```

; APPLICANT: Kekuda, Ramesh
; APPLICANT: Mezes, Peter
; APPLICANT: Millet, Isabelle
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Patturajan, Meera
; APPLICANT: Rieger, Daniel
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-462D
; CURRENT APPLICATION NUMBER: US/10/262,445
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/327,454
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/329,414
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/330,142
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/341,058
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/343,629
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 132
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-262-445-132

Query Match 41.6%; Score 47; DB 15; Length 568;
Best Local Similarity 64.3%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VWLEAVSDLHRL 22
DB 269 VQLEQISDLHKS 282

RESULT 15
US-10-042-865-38
; Sequence 38, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchiernev, Velizar T
; APPLICANT: Miller, Charles E

```

```

; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Perence L
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-042-865-38

Query Match 41.6%; Score 47; DB 15; Length 568;
Best Local Similarity 64.3%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 VWLEAVSDLHRL 22
DB 269 VQLEQISDLHKS 282

Search completed: November 11, 2004, 02:43:17
Job time : 26.4953 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 5.39623 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-61

Perfect score: 113

Sequence: 1 VVDVPDFIVWLEFAVSDLHRL 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.5	48.2	349	2 T25348	hypothetical prote
2	53	46.9	806	2 H70647	probable NADH2 den
3	51	45.1	349	2 T26367	hypothetical prote
4	50	44.2	294	2 AG0913	probable adenine-s
5	50	44.2	363	2 AG3002	3-dehydroquinat s
6	50	44.2	377	2 A93281	3-dehydroquinat s
7	48.5	42.9	328	2 F83599	hypothetical prote
8	48	42.5	82	2 AB0272	hypothetical prote
9	47	41.6	292	2 A84487	modification methy
10	47	41.6	557	2 JC6539	tumor necrosis fac
11	46	40.7	197	2 A87286	conserved hypotet
12	46	40.7	296	2 F91145	probable methyltra
13	46	40.7	296	2 B85991	probable methyltra
14	46	40.7	296	2 H51118	hypothetical adeni
15	46	40.7	378	2 A32257	3-dehydroquinat s
16	46	40.7	749	2 C84508	probable cap-bindi
17	46	40.7	829	2 T45036	hypothetical prote
18	46	40.7	3678	2 S28916	dystrophin - mouse
19	46	40.7	3685	1 A37605	dystrophin, muscle
20	45	39.8	163	2 B84308	hypothetical prote
21	45	39.8	536	2 T41652	hypothetical prote
22	45	39.8	539	2 D36904	carbon dioxide con
23	44.5	39.4	109	2 S38836	hypothetical prote
24	44.5	39.4	317	2 G75041	ornithine carbamoy
25	44.5	39.4	317	2 G71119	probable ornithine
26	44	38.9	119	2 T22181	hypothetical prote
27	44	38.9	124	2 B71494	probable dihydron
28	44	38.9	129	2 F81173	hypothetical prote
29	44	38.9	129	2 D81931	hypothetical prote

30 44 38.9 139 2 E81173
31 44 38.9 202 2 AF2831
32 44 38.9 202 2 B97609
33 44 38.9 207 2 C69338
34 44 38.9 271 2 C47127
35 44 38.9 273 2 A32123
36 44 38.9 274 2 JN0908
37 44 38.9 386 2 AE1853
38 44 38.9 441 2 T49113
39 44 38.9 493 2 T22180
40 44 38.9 563 2 F87986
41 44 38.9 1140 2 T41457
42 44 38.9 1203 2 T04294
43 44 38.9 1489 2 G71406
44 44 38.9 1633 2 T01879
45 43.5 38.5 467 1 C69301

ALIGNMENTS

RESULT 1

T25348

hypothetical protein T27C5.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T25348

R;Cummings, P.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z20020

A;Accession: T25348

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-349 <MIL>

A;Cross-references: UNIPROT:O45849; EMBL:Z82058; PIDN:CAB04866.1; GSPDB:GN000023; CBSP:T

A;Experimental source: clone T27C5

C;Genetics:

A;Gene: CESP.T27C5.5

A;Map position: 5

A;Introns: 58/2; 171/3; 273/1

C;Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 48.2%; Score 54.5; DB 2; Length 349;

Best Local Similarity 43.5%; Pred. No. 2.3;

Matches 10; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

Oy 1 VVDVPDFIVWLEEE-AVSDLHRL 22

Db 180 IIDDPDFHWEPEPMYAIHYAI 202

RESULT 2

H70647

probable NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain G - Mycobacterium tuberc

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C;Accession: H70647

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: H70647

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-806 <COL>

A;Cross-references: UNIPROT:P95175; GB:Z83867; GB:AL123456; NID:G3261695; PIDN:CAB06281

A;Experimental source: strain H37Rv

C;Genetics:

A;Gene: nuoG

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 11

C;Keywords: NAD; oxidoreductase

Query Match 46.9%; Score 53; DB 2; Length 806;
Best Local Similarity 50.0%; Pred. No. 9.3;
Matches 12; Conservative 3; Mismatches 7; Indels 2; Gaps 1;
QY 1 VVDVDFIVW--LEBAVSDLHRL 22
Db 769 VTDMDRVVWLEPLNSAGSVHRQL 792

RESULT 3
T26367
Hypothetical protein Y102A5C.31 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26367
R;Gardner, A.
Submitted to the EMBL Data Library, September 1998
A;Reference number: Z20204
A;Accession: T26367
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-349 <WIL>
A;Cross-references: UNIPROT:O9XX66; EMBL:AL031627; PIDN:CAA20969.1; CESP:Y102A5C.31
A;Experimental source: clone Y102A5C
C;Genetics:
A;Gene: CESP:Y102A5C.31
A;Introns: 58/2; 171/3; 273/1
C;Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 45.1%; Score 51; DB 2; Length 349;
Best Local Similarity 61.5%; Pred. No. 7.6;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 VVDVDFIVWLEE 13
Db 180 IIDDVDFVWLEE 192

RESULT 4
AG0913
probable adenine-specific DNA-modification methylase Srv3566 [imported] - Salmonella enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 22-Jun-2003
C;Accession: AG0913
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Cronin, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
C;Accession: AG0913
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-294 <PAR>
A;Cross-references: GB:ALJ513382; PIDN:CAD07901.1; PID:G16504448; GSPDB:GN00176
C;Genetics:
A;Gene: Srv3566
C;Superfamily: type II site-specific DNA-methyltransferase

Query Match 44.2%; Score 50; DB 2; Length 294;
Best Local Similarity 40.0%; Pred. No. 8.9;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 3 DVVDFIVWLEAVSDLHRL 22
Db 57 DEASFLAWLYECIDCHRYL 76
RESULT 5

AG3002
3-dehydroquininate synthase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AG3002
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, E.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, S.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AG3002
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-363 <KUR>
A;Cross-references: GB:AE008689; PIDN:RAL44437.1; PID:G17742038; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: aroB
A;Map position: linear chromosome
C;Superfamily: 3-dehydroquininate synthase; 3-dehydroquininate synthase homology

Query Match 44.2%; Score 50; DB 2; Length 363;
Best Local Similarity 44.4%; Pred. No. 11;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 1 VVDVDFIVWLEAVSDL 18
Db 186 LIDKPDFFWLEKNWDDI 203

RESULT 6
A99281
3-dehydroquininate synthase [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: A99281
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A99281
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-377 <KUR>
A;Cross-references: UNIPROT:Q8U9V0; GB:AE007870; PIDN:AAK89771.1; PID:G15159695; GSPDB: A99281
C;Genetics:
A;Gene: AGR L 2401
A;Map position: linear chromosome
C;Superfamily: 3-dehydroquininate synthase; 3-dehydroquininate synthase homology

Query Match 44.2%; Score 50; DB 2; Length 377;
Best Local Similarity 44.4%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 1 VVDVDFIVWLEAVSDL 18
Db 200 LIDKPDFFWLEKNWDDI 217

RESULT 7
F83599
hypothetical protein PA0360 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83599
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

C:Accession: F91145
 R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: F91145
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-236 <HAY>
 A:Cross-references: UNIPROT:Q8X9A6; GB:BA000007; PIDN:BA37557.1; PID:gl13363607; GSPDB:G
 A:Experimental source: strain O157:H7, substrain RMD 050952
 C:Genetics:
 A:Gene: ECs4134
 C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 40.7%; Score 46; DB 2; Length 296;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 7 FIVWLEAVSDLHRL 22
 ||||| :|||
 Db 63 FIDWLEVIACHRVL 78

RESULT 13

B85991
 probable methyltransferase yhdJ [imported] - *Escherichia coli* (strain O157:H7, substrain
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: B85991
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: B85991
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-296 <STO>
 A:Cross-references: UNIPROT:Q8X9A6; GB:AE005174; NID:gl2517892; PIDN:AAG58390.1; GSPDB:G
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: yhdJ
 C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 40.7%; Score 46; DB 2; Length 296;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 7 FIVWLEAVSDLHRL 22
 ||||| :|||
 Db 63 FIDWLEVIACHRVL 78

RESULT 14

H65118
 hypothetical adenine-specific methylase in *fts*-envR intergenic region - *Escherichia coli*
 C:Species: *Escherichia coli*
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 A:Accession: H65118; D47043; C28207
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: H65118
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-236 <BLAT>
 A:Cross-references: UNIPROT:P28638; GB:AE000405; GB:U00096; NID:gl789659; PIDN:AAC76294.
 A:Experimental source: strain K-12, substrain MG1655
 R:Ball, C.A.; Osuna, R.; Ferguson, K.C.; Johnson, R.C.

J. Bacteriol. 174, 8043-8056, 1992
 A:Title: Dramatic changes in *Fis* levels upon nutrient upshift in *Escherichia coli*.
 A:Reference number: A47043; MUID:93094136; PMID:1459953
 A:Accession: D47043
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 3-29, 'N', 31-102 <BAL>
 A:Cross-references: GB:M95784; NID:gl45974; PIDN:AAA23784.1; PID:gl145978
 A:Note: Sequence extracted from NCBI backbone (NCBIN:119972, NCBI:119976)
 R:Johnson, R.C.; Ball, C.A.; Pfeiffer, D.; Simon, M.I.
 Proc. Natl. Acad. Sci. U.S.A. 85, 3484-3488, 1988
 A:Title: Isolation of the gene encoding the *Hin* recombinational enhancer binding protein
 A:Reference number: A28207; MUID:88217925; PMID:2835774
 A:Accession: C28207
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 3-13 <JOH>
 A:Cross-references: GB:J03245; NID:gl45970; PIDN:AAA83857.1; PID:gl128954
 C:Genetics:
 A:Gene: yhdJ
 C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 40.7%; Score 46; DB 2; Length 296;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 7 FIVWLEAVSDLHRL 22
 ||||| :|||
 Db 63 FIDWLEVIACHRVL 78

RESULT 15

AF3257
 3-dehydroquininate synthase (EC 4.2.3.4) [imported] - *Brucella melitensis* (strain 16M)
 C:Species: *Brucella melitensis*
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 C:Accession: AF3257
 R:DeVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mue, C.; Los, T.; Ivanova,
 ; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella meliten*
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AF3257
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-378 <KUR>
 A:Cross-references: UNIPROT:Q8YJN9; GB:AE008917; PIDN:AAL51225.1; PID:gl7981913; GSPDB:
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI0043
 A:Map position: 1

Query Match 40.7%; Score 46; DB 2; Length 378;
 Best Local Similarity 58.3%; Pred. No. 45;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VDVDPDFIWLLE 12
 :||| :|||
 Db 200 LIDRDPFFAWLE 211

Search completed: November 10, 2004, 14:52:28
 Job time : 6.39623 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 29.3057 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-61
Perfect score: 113
Sequence: 1 VVDVDPFIVMLEAEVSDLHRL 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	22	Q96JA7	Q96ja7 homo sapien
2	54.5	48.2	349	O45849	O45849 caenorhabdi
3	54	47.8	357	Q7Y3V5	Q7Y3V5 bacterioph
4	53	46.9	806	1 NUOG MYCBO	P59662 mycobacteri
5	53	46.9	806	1 NUOG MYCTU	P95175 mycobacteri
6	52	46.0	3094	2 Q8UDF3	Q8udf3 plasmodium
7	52	46.0	3094	2 AAQ73930	AAq73930 plasmodiu
8	51	45.1	349	2 Q9XX66	Q9xx66 caenorhabdi
9	51	45.1	894	2 Q7PZJ5	Q7pzj5 anopheles g
10	50	44.2	227	2 Q6RFG9	Q6rfg9 pycnococcus
11	50	44.2	227	2 AAR99499	Aar99499 pycnococc
12	50	44.2	294	2 Q8XF74	Q8xf74 salmoneilla
13	50	44.2	294	2 Q7CPL9	Q7cpl9 salmoneilla
14	50	44.2	377	1 AROB-AGRT5	Q8u9v0 agrobacteri
15	50	44.2	395	2 Q728W5	Q728w5 desulfovibri
16	50	44.2	395	2 AAS96959	Aas96959 desulfovi
17	49	43.4	389	2 Q9LW88	Q9lwe8 oryza sativ
18	49	43.4	3680	1 DMD CANFA	O97592 canis famil
19	48.5	42.9	328	2 Q31604	Q3i604 pseudomonas
20	48.5	42.9	513	2 Q9W067	Q9w067 drosophila
21	48.5	42.9	588	2 Q86IH2	Q86ih2 dictyosteli
22	48	42.5	82	2 Q8ZE32	Q8ze32 yersinia pe
23	48	42.5	82	2 AAS62246	Aas62246 yersinia
24	48	42.5	345	2 Q91UW7	Q91uw7 uncultured
25	48	42.5	610	1 GLMS_PBSRM	O87tt8 p glucosami
26	48	42.5	856	2 Q7S1G9	Q7s1g9 neurospora
27	48	42.5	1653	2 Q8UDW9	Q8udw9 plasmodium
28	48	42.5	1653	2 AAQ73924	AAq73924 plasmodiu
29	48	42.5	3056	2 Q8I639	Q8i639 plasmodium
30	48	42.5	3064	2 Q6UDW7	Q6udw7 plasmodium
31	48	42.5	3064	2 AAQ73926	AAq73926 plasmodiu

32	47.5	42.0	485	2	Q7Q6Q6	Q7q6q6 anopheles g
33	47	41.6	292	1	MTWS_METJA	Q58993 methanococc
34	47	41.6	429	2	Q6PB16	Q6pb16 xenopus lae
35	47	41.6	429	2	AAS59968	Aas59968 xenopus l
36	47	41.6	557	1	TRA5_HUMAN	O00463 homo sapien
37	47	41.6	1022	2	Q82MA7	Q82ma7 streptomyce
38	47	41.6	1234	2	Q8VD34	Q8vd34 mus musculu
39	46.5	41.2	2766	1	THYG_MOUSE	O08710 mus musculu
40	46	40.7	143	2	Q688N6	Q688n6 plasmodium
41	46	40.7	143	2	AAR32054	Aar32054 plasmodiu
42	46	40.7	166	2	Q89XN3	Q89xn3 bradyrhizob
43	46	40.7	197	2	Q9ABD1	Q9abd1 caulobacter
44	46	40.7	238	2	Q61R98	Q61r98 xenopus lae
45	46	40.7	238	2	AAH71002	Aah71002 xenopus l

ALIGNMENTS

RESULT 1

Q96JA7	PRELIMINARY;	PRT;	22 AA.
ID	Q96JA7	PRELIMINARY;	PRT;
AC	Q96JA7;		
DT	01-DEC-2001 (Tremblrel. 19, Created)		
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)		
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)		
DE	Bcl-XL-binding protein v18 (Fragment).		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Sutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Kidney;		
RA	MEDLINE=21293069; PubMed=11283018;		
RT	Hammond P.W., Alpin J., Rise C.E., Wright M., Kreider B.L.;		
RT	"In Vitro Selection and Characterization of Bcl-XL-binding Proteins		
RT	from a Mix of Tissue-specific mRNA Display Libraries.";		
RL	J. Biol. Chem. 276:20898-20906(2001).		
DR	EMBL; AF357526; AAK60630.1; -.		
FT	NON_TER	1	
FT	NON_TER	22	
SQ	SEQUENCE	22 AA;	2523 MW; 0F87448F26BCF1F9 CRC64;

Query Match 100.0%; Score 113; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VVDVDPFIVMLEAEVSDLHRL 22
Db	1	VVDVDPFIVMLEAEVSDLHRL 22

RESULT 2

C45849	PRELIMINARY;	PRT;	349 AA.
ID	C45849	PRELIMINARY;	PRT;
AC	C45849;		
DT	01-JUN-1998 (Tremblrel. 06, Created)		
DT	01-JUN-1998 (Tremblrel. 06, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	C. elegans SRH-132 protein. (Corresponding sequence T27C5.5).		
GN	Names=srh-132; ORNames=127C5.5;		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;		
OC	Rhabditidae; Pelodierinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RA	MEDLINE=99069613; PubMed=9851916;		
RT	none;		
RT	"Genome sequence of the nematode C.elegans: A platform for		
RT	investigating biology.";		

```

RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Cummings PN.;
DR EMBL; 282058; CAS04866.1; -
DR PIR; T25348; T25348.
DR WormPep; T27C5.5; CE16499.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
DR InterPro; IPR003003; 7TM chemerecept.
DR InterPro; IPR000168; Nm7TM_chemerecept.
DR Pfam; PF01604; 7tm_5; 1
SQ SEQUENCE 349 AA; 40102 MW; DB40D702EC098901 CRC64;

Query Match 48.2%; Score 54.5; DB 2; Length 349;
Best Local Similarity 43.5%; Pred. No. 7.9;
Matches 10; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

QY 1 VVDPDFVWLEE-AVSDLHRL 22
DB 180 IIDDPDFHLEWEPYAIHYAI 202

RESULT 3
QY3V5 PRELIMINARY; PRT; 357 AA.
AC QY3V5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DNA adenine-methylase.
OS Bacteriophage PY54.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=172667;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=1289832;
RA Hertwig S., Klein I., Schmidt V., Beck S., Hammerl J.A., Appel B.;
RT "Sequence analysis of the genome of the temperate Versinia
enterocolitica phage PY54."
RL J. Mol. Biol. 331:605-622 (2003).
DR EMBL; AJ564013; CAD91820.1; -
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; F:DNA methylation; IEA.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Methyltransferase.
SQ SEQUENCE 357 AA; 40867 MW; 2152FA565D63927 CRC64;

Query Match 47.8%; Score 54; DB 2; Length 357;
Best Local Similarity 47.4%; Pred. No. 9.6;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 VPDFFVWLEE-AVSDLHRL 22
DB 53 VPDFFVWLEE-AVSDLHRL 71

RESULT 4
NUOG_MYCBO STANDARD; PRT; 806 AA.
AC P59962;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE NADH-quinone oxidoreductase chain G (EC 1.6.99.5) (NADH dehydrogenase
I, chain G) (NDH-1, chain G).
GN Name=nuoG; OrderedLocusNames=M53175;

OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Mensemer C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
CC -1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
sulfur (Fe-S) centers, to quinones in the respiratory chain. The
immediate electron acceptor for the enzyme in this species is
believed to be menaquinone. Couples the redox reaction to proton
translocation (for every two electrons transferred, four hydrogen
ions are translocated across the cytoplasmic membrane), and thus
conserves the redox energy in a proton gradient (By similarity).
CC -1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -1- COFACTOR: Binds 1 2Fe-2S cluster and 3 4Fe-4S clusters per subunit
(by similarity).
CC -1- SIMILARITY: Belongs to the complex I 75 kDa subunit family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: BX248345; CAD95267.1; -
CC InterPro; IPR000283; Complex1_75K.
CC InterPro; IPR001041; Ferredoxin.
CC InterPro; IPR006656; Molybdopterin.
CC InterPro; IPR006963; Molybdop_Fe4S4.
CC InterPro; IPR006657; Mol_dinuc_bind.
CC InterPro; IPR010228; NuOG.
CC Pfam; PF00111; Fer2; 1.
CC Pfam; PF00384; Molybdopterin; 1.
CC Pfam; PF04879; Molybdop_Fe4S4; 1.
CC Pfam; PF01588; Molybdop_Binding; 1.
CC TIGRfams; TIGR01973; NuOG; 1.
CC PROSITE; PS00641; COMPLEX1_75K_1; 1.
CC PROSITE; PS00642; COMPLEX1_75K_2; 1.
CC PROSITE; PS00643; COMPLEX1_75K_3; 1.
CC 2Fe-2S; 4Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding;
KW NAD; Oxidoreductase; Quinone.
FT METAL 49 49 Iron-sulfur 1 (2Fe-2S) (By similarity).
FT METAL 60 60 Iron-sulfur 1 (2Fe-2S) (By similarity).
FT METAL 63 63 Iron-sulfur 1 (2Fe-2S) (By similarity).
FT METAL 77 77 Iron-sulfur 1 (2Fe-2S) (By similarity).
FT METAL 111 111 Iron-sulfur 2 (4Fe-4S) (By similarity).
FT METAL 115 115 Iron-sulfur 2 (4Fe-4S) (By similarity).
FT METAL 118 118 Iron-sulfur 2 (4Fe-4S) (By similarity).
FT METAL 124 124 Iron-sulfur 2 (4Fe-4S) (By similarity).
FT METAL 164 164 Iron-sulfur 3 (4Fe-4S) (By similarity).
FT METAL 167 167 Iron-sulfur 3 (4Fe-4S) (By similarity).
FT METAL 170 170 Iron-sulfur 3 (4Fe-4S) (By similarity).
FT METAL 214 214 Iron-sulfur 3 (4Fe-4S) (By similarity).
FT METAL 240 240 Iron-sulfur 4 (4Fe-4S) (Potential).
FT METAL 243 243 Iron-sulfur 4 (4Fe-4S) (Potential).
FT METAL 247 247 Iron-sulfur 4 (4Fe-4S) (Potential).
FT METAL 275 275 Iron-sulfur 4 (4Fe-4S) (Potential).
SQ SEQUENCE 806 AA; 85412 MW; D674D74FA65FF3B81 CRC64;

Query Match 46.9%; Score 53; DB 1; Length 806;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 12; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

```

QY 1 VDVDPFIVW--LEEAVSDLHRL 22
 DB 769 VTDMPDRVWLPNSAGSVHRQL 792

RESULT 5
 NUOG MYCTU
 ID NUOG MYCTU STANDARD; PRT; 806 AA.
 AC P95175;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE NADH-quinone oxidoreductase chain G (EC 1.6.99.5) (NADH dehydrogenase
 DE I, chain G) (NDH-1, chain G).
 GN Name=nuoG; OrderedLocNames=rv3151, MT3239; ORFNames=MTYC03A2.07c;
 OS Mycobacterium tuberculosis
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 EX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/311159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
 RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
 RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
 RA Sultson J.E., Taylor K., Whitthead S., Barrall B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RN Nature 393:537-544 (1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 EX MEDLINE=22208494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
 RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490 (2002).
 CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
 CC sulfur (Fe-S) centers, to quinones in the respiratory chain. The
 CC immediate electron acceptor for the enzyme in this species is
 CC believed to be menaquinone. Couples the redox reaction to proton
 CC translocation (for every two electrons transferred, four hydrogen
 CC ions are translocated across the cytoplasmic membrane), and thus
 CC conserves the redox energy in a proton gradient (By similarity).
 CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
 CC -!- COFACTOR: Binds 1 2Fe-2S cluster and 3 4Fe-4S clusters per subunit
 CC (By similarity).
 CC -!- SIMILARITY: Belongs to the complex I 75 kDa subunit family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; BX842582; CAB06288.1; -
 DR EMBL; AE007138; AAK47578.1; ALT_INIT.
 DR PIR; H70647; H70647.
 DR TIGR; MT3239; -
 DR InterPro; IPR001450; 4Fe4S_ferredoxin.

DR InterPro; IPR000283; Complex1_75K.
 DR InterPro; IPR001041; Ferredoxin.
 DR InterPro; IPR006656; Molybdopterin.
 DR InterPro; IPR006963; Molybdop_Fe4S4.
 DR InterPro; IPR006657; Mol_dinuc_bind.
 DR InterPro; IPR010228; NuOG.
 DR Pfam; PF00111; Fer2; 1.
 DR Pfam; PF00384; Molybdopterin; 1.
 DR Pfam; PF04879; Molybdop_Fe4S4; 1.
 DR Pfam; PF01568; Molybdop_Binding; 1.
 DR Pfam; PF00353; 4Fe4SFRDXIN.
 DR TIGRFAMS; TIGR01973; NuOG; 1.
 DR PROSITE; PS00641; COMPLEX1_75K_1; 1.
 DR PROSITE; PS00642; COMPLEX1_75K_2; 1.
 DR PROSITE; PS00643; COMPLEX1_75K_3; 1.
 KW 2Fe-2S; 4Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding;
 KW NAD; Oxidoreductase; Quinone.
 FT METAL 49 49 Iron-sulfur 1 (2Fe-2S) (By similarity).
 FT METAL 60 60 Iron-sulfur 1 (2Fe-2S) (By similarity).
 FT METAL 63 63 Iron-sulfur 1 (2Fe-2S) (By similarity).
 FT METAL 77 77 Iron-sulfur 1 (2Fe-2S) (By similarity).
 FT METAL 111 111 Iron-sulfur 2 (4Fe-4S) (By similarity).
 FT METAL 115 115 Iron-sulfur 2 (4Fe-4S) (By similarity).
 FT METAL 118 118 Iron-sulfur 2 (4Fe-4S) (By similarity).
 FT METAL 124 124 Iron-sulfur 2 (4Fe-4S) (By similarity).
 FT METAL 164 164 Iron-sulfur 3 (4Fe-4S) (By similarity).
 FT METAL 167 167 Iron-sulfur 3 (4Fe-4S) (By similarity).
 FT METAL 170 170 Iron-sulfur 3 (4Fe-4S) (By similarity).
 FT METAL 214 214 Iron-sulfur 3 (4Fe-4S) (By similarity).
 FT METAL 240 240 Iron-sulfur 4 (4Fe-4S) (Potential).
 FT METAL 243 243 Iron-sulfur 4 (4Fe-4S) (Potential).
 FT METAL 247 247 Iron-sulfur 4 (4Fe-4S) (Potential).
 FT METAL 275 275 Iron-sulfur 4 (4Fe-4S) (Potential).
 FT CONFLICT 474 474 I -> M (in Ref. 2).
 FT CONFLICT 604 604 T -> A (in Ref. 2).
 SQ SEQUENCE 806 AA; 85423 MW; 519A1EA83181064 CRC64;

Query Match 46.9%; Score 53; DB 1; Length 806;
 Best Local Similarity 50.0%; Pred. NO. 32;
 Matches 12; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 1 VDVDPFIVW--LEEAVSDLHRL 22
 DB 769 VTDMPDRVWLPNSAGSVHRQL 792

RESULT 6
 Q6UDW3
 ID Q6UDW3 PRELIMINARY; PRT; 3094 AA.
 AC Q6UDW3;
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Erythrocyte membrane protein 1.
 GN Name=var;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malayan-Camp;
 RX PubMed=14651636;
 RA Kraemer S.M., Smith J.D.;
 RT "Evidence for the importance of genetic structuring to the structural
 RT and functional specialization of the Plasmodium falciparum var gene
 RT family."
 RL Mol. Microbiol. 50:1527-1538 (2003).
 DR EMBL; AY372127; AAQ73930.1; -
 DR InterPro; IPR000875; Cecropin.
 DR PROSITE; PS00268; CECROPIN; UNKNOWN 1.
 SQ SEQUENCE 3094 AA; 359444 MW; 31C0DA3CB5C17A07 CRC64;

Query Match 46.0%; Score 52; DB 2; Length 3094;

```

Best Local Similarity 52.9%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 VPDFIVWLEEAUSDHRA 20
Db 273 VPQFLRWLTWIEDYLR 289

RESULT 7
AAQ73930
ID AAQ73930 PRELIMINARY; PRT; 3094 AA.
AC AAQ73930;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte membrane protein 1.
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malayan-Camp;
RX PubMed=14651636;
RA Kraemer S.M., Smith J.D.;
RT "Evidence for the importance of genetic structuring to the structural
RT and functional specialization of the Plasmodium falciparum var gene
RT family."
RL Mol. Microbiol. 50:1527-1538(2003).
RL EMBL; AY372127; AAQ73930.1; -.
SQ SEQUENCE 3094 AA; 35944 MW; 31C0DA3CB5C17A07 CRC64;

Query Match 46.0%; Score 52; DB 2; Length 3094;
Best Local Similarity 52.9%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 VPDFIVWLEEAUSDHRA 20
Db 273 VPQFLRWLTWIEDYLR 289

RESULT 8
Q9XX66
ID Q9XX66 PRELIMINARY; PRT; 349 AA.
AC Q9XX66;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Y102A5C.31.
GN Name=Y102A5C.31;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Gardner A.E.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBSJ databases.
RL EMBL; AL031627; CAA20969.1; -.
DR F1R; T26367; T26367.
DR WormPep; Y102A5C.31; CE20402.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
DR InterPro; IPR003003; 7TM_chemrecept2.

```

```

DR InterPro; IPR000168; Nm7TM_chemrecept.
DR Pfam; PF01604; 7tm5; 1.
KW Hypothetical protein.
SQ SEQUENCE 349 AA; 40149 MW; 64664F68B6D98DD5 CRC64;

Query Match 45.1%; Score 51; DB 2; Length 349;
Best Local Similarity 61.5%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVDVPDFIVWLEE 13
Db 180 IIDDPDFHLWLEE 192

RESULT 9
Q7PZJ5
ID Q7PZJ5 PRELIMINARY; PRT; 894 AA.
AC Q7PZJ5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP9332 (Fragment).
GN Name=agCG54448; ORFNames=ENSANGG00000012809;
OS Anopheles gambiae str. PE8T.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PE8T;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008986; EAA00658.1; -.
DR InterPro; IPR003409; MORN.
DR Pfam; PF02493; MORN; 5.
FT NON TER 1
SQ SEQUENCE 894 AA; 101315 MW; D2A93539D3653B79 CRC64;

Query Match 45.1%; Score 51; DB 2; Length 894;
Best Local Similarity 47.6%; Pred. No. 73;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 VVDVPDFIVWLEEAUSDHRA 21
Db 112 VVQYPLNLVWITVSELSHS 132

RESULT 10
Q6RFG9
ID Q6RFG9 PRELIMINARY; PRT; 227 AA.
AC Q6RFG9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (fragment).
GN Name=Cox2;
OS Pycnococcus provasolii.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae; Mamiellales;
OC Pycnococaceae; Pycnococcus.
OX NCBI_TaxID=41880;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14761653;
RA Hackett J.D., Xoon H.S., Soares M.B., Bonaldo M.F., Casavant T.L.,
RA Schatz T.E., Nosenko T., Bhattacharya D.;
RT "Migration of the plastid genome to the nucleus in a peridinin
RT dinoflagellate."
RL Curr. Biol. 14:213-218(2004).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory

```


chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. Subunit 2 transfers the electrons from cytochrome c via its binuclear copper A center to the bimetallic center of the catalytic subunit 1 (By similarity).

-1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferrocycytochrome c + 2 H2O.

-1- COPACTOR: Copper A (By similarity).

-1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).

-1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family. EMBL: AY500835; AAR99499.1; -.

GO: GO:0005739; C:mitochondrion; IEA.

InterPro: IPR001505; Copper_CuA.

InterPro: IPR008972; Cupredoxin.

InterPro: IPR002429; Cyt_c_ox_2.

Pfam: PF00116; COX2; 1.

Pfam: PF02790; COX2 TM; 1.

PRINTS: PR01166; CYCOXIDASEII.

ProDom: PD000131; Copper_CuA; 1.

Copper; Electron transport; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.

NON_TER 1

SEQUENCE 227 AA, 25590 MW; CF7637C75DB2899B CRC64;

Query Match · 44.2%; Score 50; DB 2; Length 227;

Best Local Similarity 47.1%; Pred. No. 25;
Matches 8; Conservative 4; Mismatches 5; Indels

Qy 2 VDPDFIWL EEA VSDL 18

Db 207 VSFDDFIWLEAKISEI 223

RESULT 11

REC'D 1
AAR99499
ID AAR99499
PRELIMINARY: PRT: 227 AA.

AC AAR99499;
DT 29-MAR-2004 (Tremblrel. 27. Created)

[illegible]

DE Cytochrome oxidase subunit II (Fragment).

GN COX2.

OS *Pycnococcus provasolii*.

OG Mitochondrion.

OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae; Mamiellales;
OC Pycnococcaceae; Pycnococcus.

OX NCBI TaxID=41880;

RN [1]

RP SEQUENCE FROM N. A.

PubMed=14761653;
Hackett J.D., Yoon H.S., Soares M.B., Ronaldo M.F., Casavant T.L.,

RA Scheetz T.E., Nosenko T., Bhattacharya D.;
RT "Migration of the plastid genome to the nucleus in a peridinin

RT dinoflagellate.";

RL Curr. Biol. 14:213-218(2004).

DR EMBL; AY500835;
MW Mitochondrion

MISCELLANEOUS.	
KW	NON TER 1
FT	NON TER 1

SO SEQUENCE 227 AA: 25590 MW: CF7637C75DB2899B CRC64: 1

Query Match	44.28;	Score 50;	DB 2;	Length 227;
-------------	--------	-----------	-------	-------------

Best Local Similarity 47.1%; Pred. No. 25;
Matches 8; Conservative 4; Mismatches 5; Indels

Qy 2 VDVPDFIVWLEEA VSDL 18

..

Db 207 VSFDDFIWLEAKISEI 223

RESULT 12

Q8XF74
ID O8XF74
PRELIMINARY: PRT: 294 AA.

AC DT DT DT DE GN OS OC OC OX RN RP RC RX RA RA RT RT RL RN RP RC RX RA RA RA RA RA RA RA RT RT RL DR DR DR DR DR DR DR DR DR DR DR DR KW SQ

Query Match

Query Match	100%	Score 50, 52, 57	Length 11
Best Local Similarity	40.0%	Pred. No. 33;	
Matches	8: Conservative	3: Mismatches	9: Indels

3 DVPDFIVWL EEA VSD LHRAL 22
Ov
Ov

57 DEASFLAWLYECIDECRVL 766

RESULT 13

Q7CPL9

ID	Q7CPL9	PRELIMINARY;	PRT;	294 AA.
----	--------	--------------	------	---------

AC Q7CPL9;

DT 05-JUL-2004 (TREMBlrel. 27, Created)

DT	05-JUL-2004 (TREMBlRel. 27, Last sequence update)	DT	05-JUL-2004 (TREMBlRel. 27, Last annotation update)
DT	05-JUL-2004 (TREMBlRel. 27, Last sequence update)	DT	05-JUL-2004 (TREMBlRel. 27, Last annotation update)

DE Putative methyltransferase (EC 2.1.1.72).

```

DE      Name=yhdJ; OrderedLocusNames=STM3386;
GN      Functional meaning of the translated protein (NC 2.1.1)

```

OS *Salmonella typhimurium*.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; OC

OC. Enterobacteriaceae; Salmonella.

```

OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008856; AAL22255.1; -;
DR GO; GO:0008168; P: methyltransferase activity; IEA.
DR GO; GO:0009007; P: site-specific DNA-methyltransferase (adenin. .); IEA.
DR GO; GO:0016740; P: transferase activity; IEA.
DR InterPro; IPR002295; D2IN6_mtfrase.
DR InterPro; IPR001091; MetTrans_CN4.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PRINTS; PR00508; S2IN6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN1.
KW Complete proteome; Methyltransferase; Transferase.
SQ SEQUENCE 294 AA; 33554 MW; C674376A21C04285 CRC64;

Query Match 44.2%; Score 50; DB 2; Length 294;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 DVPDFIVLVEAEVSLHRL 22
DB 57 DEASFLAWLYECIDCHRVL 76

RESULT 14
AROB AGRT5
ID AROB AGRT5 STANDARD; PRT; 377 AA.
AC Q8U9V0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 3-dehydroquininate synthase (EC 4.2.3.4).
GN Name=aroB; OrderedLocusNames=atu3625, AGR_L_2401;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OC NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Sebubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,

```

```

RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
CC -1- CATALYTIC ACTIVITY: 3-deoxy-arabino-heptulosonate 7-phosphate = 3-
CC dehydroquininate + phosphate.
CC -1- COFACTOR: NAD and a divalent metal cation (By similarity).
CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC second step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the dehydroquinase synthase family.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE008291; AAL44437.1; ALT_INIT.
CC EMBL; AE008321; AAK89771.1; -.
CC PIR; A93281; A99281.
CC HSP; P07547; INVE.
CC HAMAP; MF_00110; -.
CC InterPro; IPR002658; DHQ_synthase.
CC Pfam; PF01761; DHQ_synthase; 1.
CC TIGRfams; TIGR01357; aroB; 1.
KW Aromatic amino acid biosynthesis; Complete proteome; Lyase; NAD.
SQ SEQUENCE 377 AA; 40050 MW; 5B75B42CE204EA17 CRC64;

Query Match 44.2%; Score 50; DB 1; Length 377;
Best Local Similarity 44.4%; Pred. No. 42;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VVDVPDFIVLVEAEVSDL 18
DB 200 LIDKPDFFWLEKWNDDI 217

RESULT 15
Q728W5 PRELIMINARY; PRT; 395 AA.
ID Q728W5
AC Q728W5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=DVU2487;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OC NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonel J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dmitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AS017317; AAS9659.1; -.
DR TIGR; DVU2487; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 395 AA; 43664 MW; 3B592825236EABEA CRC64;

```

Query Match 44.2%; Score 50; DB 2; Length 395;
 Best Local Similarity 50.0%; Pred. No. 45;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 7 FIVMLEEAVSDLHRAI 22
 |::|||::|:
 Db 37 FVLMLESAPEDVRRRAI 52

Search completed: November 10, 2004, 14:50:32
 Job time : 31.3057 secs

This Page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	61	28.2	218	4	US-09-259-991A-23272	Sequence 23272, A
2	61	28.2	369	4	US-09-489-039A-7324	Sequence 7324, A
3	61	28.2	1197	4	US-09-259-991A-30833	Sequence 30833, A
4	59	27.3	140	4	US-09-259-991A-19170	Sequence 19170, A
5	58	26.9	656	4	US-09-259-991A-22270	Sequence 22270, A
6	57.5	26.6	67	4	US-09-489-039A-9405	Sequence 9405, Ap
7	57	26.4	382	4	US-09-248-796A-18423	Sequence 18423, A
8	56	25.9	1226	1	US-08-280-443-2	Sequence 2, Appli
9	56	25.9	1226	1	US-08-457-459-2	Sequence 2, Appli
10	56	25.9	1226	1	US-08-555-678-2	Sequence 2, Appli
11	56	25.9	1226	5	PCT-US95-02475-2	Sequence 2, Appli
12	55	25.5	103	4	US-09-489-039A-7262	Sequence 7262, Ap
13	54.5	25.2	450	4	US-08-635-866C-188	Sequence 188, App
14	54.5	25.2	450	4	US-08-974-690C-188	Sequence 188, App
15	54.5	25.2	452	4	US-09-259-991A-20884	Sequence 20884, A
16	54	25.0	137	4	US-09-248-796A-28007	Sequence 28007, A
17	53	24.5	118	4	US-09-489-039A-11052	Sequence 11052, A
18	53	24.5	208	4	US-09-252-991A-23638	Sequence 23638, A
19	53	24.5	236	4	US-09-489-039A-9771	Sequence 9771, Ap
20	52	24.1	168	4	US-09-259-991A-19362	Sequence 19362, A
21	52	24.1	225	4	US-09-673-395A-480	Sequence 480, App
22	52	24.1	402	4	US-09-259-991A-31178	Sequence 31178, A
23	51.5	23.8	1083	4	US-09-259-991A-19224	Sequence 19224, A
24	51.5	23.8	408	4	US-09-259-991A-19329	Sequence 19329, A
25	51	23.6	169	4	US-09-489-039A-9755	Sequence 9755, Ap
26	51	23.6	309	4	US-09-270-767-33777	Sequence 33777, A
27	51	23.6	309	4	US-09-270-767-48994	Sequence 48994, A

Query Match 28.2%; Score 61; DB 4; Length 369;
 Best Local Similarity 75.0%; Pred. No. 1.2;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 26 GWQMPNCRSRG 37
 DB 14 GWRUPCRSRG 25

RESULT 3
 US-09-252-991A-30833
 ; Sequence 30833, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 30833
 ; LENGTH: 1197
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (1028)
 ; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

US-09-252-991A-30833

Query Match 28.2%; Score 61; DB 4; Length 1197;
 Best Local Similarity 42.5%; Pred. No. 4.6;
 Matches 17; Conservative 7; Mismatches 12; Indels 4; Gaps 2;

QY 2 RGNEFQRLDADWLSSRQRQGWQMPNCR--SRGPG 39
 DB 908 ROPDFRLRDVAEA--ALGRAAGAGPRPRRPPVGRGPG 945

RESULT 4
 US-09-252-991A-19170
 ; Sequence 19170, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 19170
 ; LENGTH: 140
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-19170

Query Match 27.3%; Score 59; DB 4; Length 140;
 Best Local Similarity 53.6%; Pred. No. 0.74;
 Matches 15; Conservative 4; Mismatches 5; Indels 4; Gaps 2;

QY 6 EF-QRLD---ADAWLSSRQRQGWQ 29

DB 47 BFGELDLRPSDEADLASREERQWQL 74

RESULT 5
 US-09-252-991A-22270
 ; Sequence 22270, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 22270
 ; LENGTH: 656
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22270

Query Match 26.9%; Score 58; DB 4; Length 656;
 Best Local Similarity 57.9%; Pred. No. 6.2;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 21 RSRQRQWQMPNCRSRGPG 39
 DB 138 RGRHAGCPFGCRSRGPG 156

RESULT 6
 US-09-489-039A-9405
 ; Sequence 9405, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 9405
 ; LENGTH: 67
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-9405

Query Match 26.6%; Score 57.5; DB 4; Length 67;
 Best Local Similarity 42.3%; Pred. No. 0.51;
 Matches 11; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

QY 16 WDLSSRSRORGW----QMPNCRSR 36
 DB 3 WSANSSASRSRWTVRAVMPSCRSR 28

RESULT 7
 US-09-248-796A-18423
 ; Sequence 18423, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstein et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A

```
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18423
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18423

Query Match      26.4%; Score 57; DB 4; Length 382;
Best Local Similarity 45.8%; Pred. No. 4.6;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 8 QLRDLADLADLSSRSRQGWQMPN 31
DB 254 QYDLAITWSLNTSKSLFWQKN 277

RESULT 8
US-08-280-443-2
; Sequence 2, Application US/08280443
; Patent No. 5643778
; GENERAL INFORMATION:
; APPLICANT: Nishikura, Kazuko
; TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,443
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-457-459-2

Query Match      25.9%; Score 56; DB 1; Length 1226;
Best Local Similarity 36.1%; Pred. No. 25;
Matches 13; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 3 RGNFQLRLADLADLSSRSRQGWQMPNCRSRGP 38
DB 84 RGRQVDIGVPRGVHLSQGLQRFQHPSPGRSLP 119

RESULT 10
US-08-555-678-2
; Sequence 2, Application US/08555678
; Patent No. 5763174
; GENERAL INFORMATION:
; APPLICANT: Nishikura, Kazuko
; TITLE OF INVENTION: RNA Editing Enzyme and Methods
; TITLE OF INVENTION: of Use Thereof
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,678
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,459
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49DUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-02275-2

Query Match 25.9%; Score 56; DB 1; Length 1226;
Best Local Similarity 36.1%; Pred. No. 25;
Matches 13; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 3 RGNFQLRLADAWDLSSRSRQGMPCNCRSRGP 38
Db 84 RGRQVDIRGVRGVHLGSQGLQRFQHPSPGRSLP 119

RESULT 12
US-09-489-039A-7262
; Sequence 7262, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7262
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7262

Query Match 25.5%; Score 55; DB 4; Length 103;
Best Local Similarity 54.5%; Pred. No. 1.9;
Matches 12; Conservative 3; Mismatches 3; Indels 4; Gaps 2;

QY 16 WDLSSRSRQGMPCNCRSRRG 37
Db 22 WAISNRS---GWRMP-CASRSG 39

RESULT 13
US-08-635-886C-188
; Sequence 188, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent version 3.1
; SEQ ID NO 188
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
```


US-08-635-886C-188

Query Match 25.2%; Score 54.5; DB 4; Length 450;
Best Local Similarity 36.8%; Pred. No. 13;
Matches 14; Conservative 5; Mismatches 16; Indels 3; Gaps 1;

Qy 2 RRGNEFQRLDADWDLSSRSRQGRQWQMPNCRSRRPG 39
Db 39 RRGPRGLGVRATRK--SERSQPRGWRQPIPKARPEG 73

RESULT 14

US-08-974-690C-188
; Sequence 188, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 188
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-188

Query Match 25.2%; Score 54.5; DB 4; Length 450;
Best Local Similarity 36.8%; Pred. No. 13;
Matches 14; Conservative 5; Mismatches 16; Indels 3; Gaps 1;

Qy 2 RRGNEFQRLDADWDLSSRSRQGRQWQMPNCRSRRPG 39
Db 39 RRGPRGLGVRATRK--SERSQPRGWRQPIPKARPEG 73

RESULT 15

US-09-252-991A-20884
; Sequence 20884, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20884
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20884

Query Match 25.2%; Score 54.5; DB 4; Length 452;
Best Local Similarity 35.3%; Pred. No. 13;
Matches 18; Conservative 5; Mismatches 11; Indels 17; Gaps 3;

Qy 2 RRGNE---FQLRLDADWDLSSRSRQGRQWQMPNCR-----SRRG 37
Db 37 RRGPRGLGVRATRK--SERSQPRGWRQPIPKARPEG 73

Db 381 RRTSERPGISLRRIRPA--LPRTRRRWQRPACRGLPLRPSSSGRVARRG 429
Search completed: November 10, 2004, 14:55:44
Job time : 15.3255 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 45.1075 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-62

Perfect score: 216

Sequence: 1 QRRGNEFQRLDLADAWLSSRRQGWQMPNCRSRGPG 39

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216	100.0	39	14 US-10-092-750-62	Sequence 62, Appl
2	64.5	29.9	286	15 US-10-425-114-66943	Sequence 66943, A
3	64.5	29.9	391	17 US-10-425-115-336744	Sequence 336744, A
4	61.5	28.5	138	17 US-10-425-115-242951	Sequence 242951, A
5	60.5	28.0	88	9 US-09-764-877-1924	Sequence 1924, Ap
6	60.5	28.0	88	15 US-10-242-515-1924	Sequence 1924, Ap
7	60.5	28.0	141	15 US-10-425-114-67835	Sequence 67835, A
8	60.5	28.0	498	17 US-10-425-115-209098	Sequence 209098, A
9	60	27.8	76	17 US-10-425-115-225012	Sequence 225012, A
10	60	27.8	190	17 US-10-425-115-36562	Sequence 36562, A
11	58	26.9	168	14 US-10-369-493-11976	Sequence 11976, A
12	56.5	26.2	78	16 US-10-437-963-142826	Sequence 142826, A
13	56.5	26.2	551	16 US-10-437-963-110875	Sequence 110875, A

14	56.5	26.2	551	16	US-10-437-963-123723	Sequence 123723, A
15	56	25.9	95	17	US-10-425-115-366113	Sequence 366113, A
16	56	25.9	716	15	US-10-412-699B-404	Sequence 404, App
17	56	25.9	1225	14	US-10-233-553-20	Sequence 20, Appl
18	55	25.5	136	15	US-10-282-122A-53966	Sequence 53966, A
19	55	25.5	877	14	US-10-369-493-19266	Sequence 19266, A
20	55	25.5	913	16	US-10-437-963-197298	Sequence 197298, A
21	54.5	25.2	450	15	US-10-651-185-188	Sequence 188, App
22	54	25.0	160	16	US-10-437-963-175669	Sequence 175669, A
23	54	25.0	210	17	US-10-425-115-186600	Sequence 186600, A
24	54	25.0	1623	16	US-10-437-963-202524	Sequence 202524, A
25	53.5	24.8	164	16	US-10-437-963-127029	Sequence 127029, A
26	53.5	24.8	169	14	US-10-156-761-7602	Sequence 7602, Ap
27	53	24.5	91	17	US-10-425-115-342590	Sequence 342590, A
28	53	24.5	160	17	US-10-425-115-218695	Sequence 218695, A
29	53	24.5	201	16	US-10-437-963-138893	Sequence 138893, A
30	53	24.5	224	17	US-10-425-115-264161	Sequence 264161, A
31	53	24.5	231	15	US-10-282-122A-59841	Sequence 59841, A
32	52.5	24.3	99	9	US-09-864-761-48852	Sequence 48852, A
33	52.5	24.3	138	16	US-10-767-701-53818	Sequence 53818, A
34	52.5	24.3	163	17	US-10-425-115-264163	Sequence 264163, A
35	52.5	24.3	210	15	US-10-425-114-70467	Sequence 70467, A
36	52.5	24.3	442	15	US-10-425-114-43237	Sequence 43237, A
37	52.5	24.3	581	15	US-10-425-114-47996	Sequence 47996, A
38	52	24.1	117	15	US-10-425-114-47996	Sequence 47996, A
39	52	24.1	137	16	US-10-437-963-189043	Sequence 189043, A
40	52	24.1	149	16	US-10-437-963-115114	Sequence 115114, A
41	52	24.1	183	16	US-10-437-963-108992	Sequence 108992, A
42	52	24.1	221	15	US-10-425-114-49030	Sequence 49030, A
43	52	24.1	222	15	US-10-424-599-151206	Sequence 151206, A
44	52	24.1	368	15	US-10-425-114-50824	Sequence 50824, A
45	52	24.1	533	14	US-10-156-761-13934	Sequence 13934, A

ALIGNMENTS

RESULT 1
US-10-092-750-62
; Sequence 62, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-62

Query Match 100.0%; Score 216; DB 14; Length 39;
Best Local Similarity 100.0%; Pred. No. 4e-21;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRRGNEFQRLDLADAWLSSRRQGWQMPNCRSRGPG 39
DB 1 QRRGNEFQRLDLADAWLSSRRQGWQMPNCRSRGPG 39
RESULT 2
US-10-425-114-66943
; Sequence 66943, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:


```
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,085
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1924
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36)_feature
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (72)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-242-515-1924

Query Match      28.0%; Score 60.5; DB 15; Length 88;
Best Local Similarity 45.9%; Pred. No. 3;
Matches 17; Conservative 1; Mismatches 10; Indels 9; Gaps 3;

Qy 12 LADAWD-LSSRSRQRG-----WQWPN---CESRRGPG 39
Db 25 LLPRFPLSSRXRSRSGGCRWWSGFSRWQCRSRSPG 61

RESULT 7
US-10-425-114-67835
; Sequence 67835, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128

; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,085
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1924
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36)_feature
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (72)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-242-515-1924

Query Match      28.0%; Score 60.5; DB 15; Length 88;
Best Local Similarity 45.9%; Pred. No. 3;
Matches 17; Conservative 1; Mismatches 10; Indels 9; Gaps 3;

Qy 12 LADAWD-LSSRSRQRG-----WQWPN---CESRRGPG 39
Db 25 LLPRFPLSSRXRSRSGGCRWWSGFSRWQCRSRSPG 61

RESULT 7
US-10-425-114-67835
; Sequence 67835, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
```

```
; SEQ ID NO 67835
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4730-006-A5_FLI.pep
US-10-425-114-67835

Query Match      28.0%; Score 60.5; DB 15; Length 141;
Best Local Similarity 34.8%; Pred. No. 4.9;
Matches 16; Conservative 7; Mismatches 14; Indels 9; Gaps 1;

Qy 3 RGNFPLRLADAWDLSSRS-----RQGWQWPNCKSRGPG 39
Db 68 RGHQRQVRGARGGDDQZALRAHPALACRQASEPFGDCRSQSGG 113

RESULT 8
US-10-425-115-209098
; Sequence 209098, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 209098
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_122292C.1.pap
US-10-425-115-209098

Query Match      28.0%; Score 60.5; DB 17; Length 498;
Best Local Similarity 41.9%; Pred. No. 17;
Matches 13; Conservative 4; Mismatches 13; Indels 1; Gaps 1;

Qy 9 LRLADAWDLSSRS-RQGWQWPNCKSRGPG 38
Db 28 ISDIGSASVSARSVAGRWGWDAPACRHRKP 58

RESULT 9
US-10-425-115-225012
; Sequence 225012, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 225012
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(76)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
```

; OTHER INFORMATION: Clone ID: MRT4577_136802C.1.pep
US-10-425-115-225012

Query Match 27.8%; Score 60; DB 17; Length 76;
Best Local Similarity 63.2%; Pred. No. 3;
Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 18 LSSRSRQSGWQWPCNCRSR 36
DB 24 LSVRSRRWPAFCRSTR 42

RESULT 10

US-10-425-115-361562
; Sequence 361562, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 361562
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(190)
; OTHER INFORMATION: unsure at all xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_92922C.1.pep
US-10-425-115-361562

Query Match 27.8%; Score 60; DB 17; Length 190;
Best Local Similarity 50.0%; Pred. No. 7.6;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 19 SRSRQSGWQWPCNCRSRGP 38
DB 89 TTSTRSGWPAFCRSTRTP 108

RESULT 11

US-10-369-493-11976
; Sequence 11976, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Cher, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11976
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-11976

Query Match 26.9%; Score 58; DB 14; Length 168;

Best Local Similarity 44.4%; Pred. No. 12;
Matches 12; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

QY 7 FQRLDLADAWDLS----SRSRQSGWQM 29
DB 118 FNKSLADRWSEISDGDIDNLRLQWQI 144

RESULT 12

US-10-437-963-142826
; Sequence 142826, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142826
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43795C.1.pep
US-10-437-963-142826

Query Match 26.2%; Score 56.5; DB 16; Length 78;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 13; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 12 LADWDLSSRQSGWQWPCNCRSRG 37
DB 17 LASAWTSSRRRAQGC-LPECRWKDG 41

RESULT 13

US-10-437-963-110875
; Sequence 110875, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 110875
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_14902C.1.pep
US-10-437-963-110875

Query Match 26.2%; Score 56.5; DB 16; Length 551;
Best Local Similarity 37.1%; Pred. No. 64;

Search completed: November 11, 2004, 02:43:18
Job time : 46.1575 secs

Matches 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 2 REGNEFQRLDADLSSRSRQGMPCRSR 36
DB 510 RYRSSELRDQKWERATR-EDRAWDFARPR 543

RESULT 14

US-10-437-963-123723
; Sequence 123723, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 123723
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2652C.1.pap
US-10-437-963-123723

Query Match 26.2%; Score 56.5; DB 16; Length 551;
Best Local Similarity 39.3%; Pred. No. 64;
Matches 11; Conservative 4; Mismatches 10; Indels 3; Gaps 1;

QY 10 RDLADAWLSSRSRQGMPCRSR 34
DB 20 REAAAWQFAGNRRRRGWGMLRPRCAS 47

RESULT 15

US-10-425-115-366113
; Sequence 366113, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 366113
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97060C.1.pap
US-10-425-115-366113

Query Match 25.9%; Score 56; DB 17; Length 95;
Best Local Similarity 35.3%; Pred. No. 13;
Matches 12; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

QY 4 GNEFQRLDADLSSRSRQGMPCRSR 37
DB 47 GEQWQXKYLGAICSCCTCGGQGWGMLRPR 80

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 9.56604 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-62

Perfect score: 216

Sequence: 1 QRGNEFQLRLDADMLSSRSRQRGMPCNCRSRGPG 39

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	29.6	466	1 B43332	glutamate decarboxylase
2	64	29.6	466	1 S24324	glutamate decarboxylase
3	64	29.6	466	2 C85726	glutamate decarboxylase
4	64	29.6	466	2 B90891	glutamate decarboxylase
5	60	27.8	466	2 E91178	glutamate decarboxylase
6	60	27.8	466	2 F86024	glutamate decarboxylase
7	59	27.3	134	2 E82953	conserved hypother
8	57	26.4	462	2 AH1130	glutamate decarboxylase
9	57	26.4	464	2 AC1370	glutamate decarboxylase
10	57	26.4	464	2 AB1740	glutamate decarboxylase
11	56	25.9	739	2 H85245	Vpi like protein [
12	56	25.9	739	2 T05163	hypothetical prote
13	56	25.9	1226	1 S65593	adenosine deaminas
14	56	25.9	2034	2 T22147	hypothetical prote
15	55	25.5	304	2 AE3623	glutamate decarboxylase
16	55	25.5	3305	2 T18358	apolipoprotein prec
17	54.5	25.2	855	2 T36132	hypothetical prote
18	54.5	25.2	154	2 JQ1137	rRNA methyltransfe
19	54.5	25.2	513	2 FC1284	genome polyprotein
20	52	24.1	194	2 C83312	hypothetical prote
21	52	24.1	746	2 T35811	probable phosphodi
22	52	24.1	1516	2 T01055	hypothetical prote
23	51.5	23.8	438	2 AD0437	conserved hypother
24	51.5	23.8	2605	2 T18552	safiramycin Mxi syn
25	51	23.6	345	2 AG3186	hypothetical prote
26	51	23.6	365	2 AH2606	transcription regu
27	51	23.6	365	2 G97388	conserved hypother
28	50.5	23.4	332	2 F75473	conserved hypother
29	50.5	23.4	621	2 E82768	conserved hypother

30 50.5 23.4 885 2 C83441 two-component sens
31 50 23.1 299 2 T43442 hypothetical prote
32 50 23.1 375 2 AB0725 ribonuclease D [im
33 50 23.1 946 2 T31488 hypothetical prote
34 50 23.1 1377 2 E86034 rnsA protein in th
35 50 23.1 1377 2 C65159 probable glucan sy
36 50 23.1 1510 2 C84727 hypothetical prote
37 50 23.1 1556 2 D36793 hypothetical prote
38 49.5 22.9 111 2 C72714 small T antigen -
39 49.5 22.9 195 1 TVVPA hypothetical prote
40 49.5 22.9 247 2 T06721 hypothetical prote
41 49.5 22.9 271 1 T36709 middle T antigen -
42 49.5 22.9 421 1 TVVPM middle T antigen -
43 49.5 22.9 440 1 TVVPM aldehyde-ferredoxi
44 49.5 22.9 621 2 A71221 aldehyde ferredoxi
45 49.5 22.9 624 2 E75188

ALIGNMENTS

RESULT 1

B43332
glutamate decarboxylase (EC 4.1.1.15) beta - Escherichia coli (strain K-12)
N;Alternate names: L-glutamate l-carboxy-lyase
C;Species: Escherichia coli
C;Date: 10-Mar-1994 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
C;Accession: B43332; S30261; H64902
R;Smith, D.K.; Kassam, T.; Singh, B.; Elliott, J.F.
J. Bacteriol. 174, 5820-5826, 1992
A;Title: Escherichia coli has two homologous glutamate decarboxylase genes that map to
A;Reference number: A43332; MUID:92394884; PMID:1522060
A;Accession: B43332
A;Molecule type: DNA
A;Residues: 1-466 <SMI>
A;Cross-references: UNIPROT:P28302; GB:M84025; NID:G146059; PIDN:AAA23834.1; PID:G14604
R;Yoshida, T.; Ueguchi, C.; Yamada, H.; Mizuno, T.
Mol. Gen. Genet. 237, 113-122, 1993
A;Title: Function of the Escherichia coli nucleoid protein, H-NS: molecular analysis of
A;Reference number: S30261; MUID:93204884; PMID:8455549
A;Accession: S30261
A;Molecule type: protein
A;Residues: 1-4, 'LQVL', 7-15 <YOS>
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64902
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-466 <BLAT>
A;Cross-references: GSI:AE000246; GB:U000096; NID:G1787764; PIDN:AAC74566.1; PID:G178776;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: gadB
A;Map position: 78 min
C;Function:
A;Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
A;Note: in E. coli, two isoforms (alpha and beta) have been found, each encoded by a se
C;Superfamily: Escherichia coli glutamate decarboxylase
C;Keywords: carbon-carbon lyase; carboxy-lyase; homohexamer; phosphoprotein; pyridoxal
F;276/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 29.6%; Score 64; DB 1; Length 466;

Best Local Similarity 50.0%; Pred No. 0.82; Mismatches 4; Gaps 1;

Matches 14; Conservative 4; Indels 6; Indels 4; Gaps 1;

QY 7 FQLRLDADMLSSRSRQRGMPCNCRSRGPG 30

DB 380 FKLKGDGPGVTLVLSRLRLRGWQVP 407

```

G85726
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85726
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamou, K.; Apodaca
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85726
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <STO>
A;Cross-references: UNIPROT:P28302; GB:AE005174; NID:g12515184; PIDN:AAG56275.1; GSPDB:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
C;Gene: gadB
C;Superfamily: Escherichia coli glutamate decarboxylase
Query Match 29.6%; Score 64; DB 2; Length 466;
Best Local Similarity 50.0%; Pred. No. 0.82;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQRLDLAD----AWDLSSRSRQGWQMP 30
Db 380 FKLKGEDPGTYLYDLSELRRLRGWQVP 407

RESULT 4
B90891
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: B90891
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasaawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90891
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <HAY>
A;Cross-references: UNIPROT:P28302; GB:BA000007; PIDN:BA035521.1; PID:g13361564; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
C;Gene: ECs2098
C;Superfamily: Escherichia coli glutamate decarboxylase
Query Match 29.6%; Score 64; DB 2; Length 466;
Best Local Similarity 50.0%; Pred. No. 0.82;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQRLDLAD----AWDLSSRSRQGWQMP 30
Db 380 FKLKGEDPGTYLYDLSELRRLRGWQVP 407

RESULT 5
E91178
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: E91178
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasaawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E91178
A;Status: preliminary
A;Molecule type: DNA

```

A;Residues: 1-466 <HAY>
A;Cross-references: UNIPROT:P58228; GB:BAQ00007; PIDN:BAB37820.1; PID:G13363871; GSPDB:G
A;Experimental source: strain O157:H7, substrain RMD 050952
C;Genetics:
A;Gene: EC64397
C;Superfamily: Escherichia coli glutamate decarboxylase

Query Match 27.8%; Score 60; DB 2; Length 466;
Best Local Similarity 46.4%; Pred. No. 2.8;
Matches 13; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLAD----AWDLSSRSRQGWQMP 30
DB 380 FKLKEGEDPGVTLYDLSERLRLGWQVP 407

RESULT 6
F86024
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: F86024
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoumis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: AB5480; MUID:21074935; PMID:11206551
A;Accession: F86024
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <STO>
A;Cross-references: UNIPROT:P58228; GB:ABQ05174; NID:G12518229; PIDN:AKG58658.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EBL933
C;Genetics:
A;Gene: gadA
C;Superfamily: Escherichia coli glutamate decarboxylase

Query Match 27.8%; Score 60; DB 2; Length 466;
Best Local Similarity 46.4%; Pred. No. 2.8;
Matches 13; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLAD----AWDLSSRSRQGWQMP 30
DB 380 FKLKEGEDPGVTLYDLSERLRLGWQVP 407

RESULT 7
B82953
conserved hypothetical protein PA5536 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: B82953
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: B82950; MUID:20437337; PMID:10984043
A;Accession: B82953
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-134 <STO>
A;Cross-references: UNIPROT:Q9HT38; GB:ABQ04966; GB:ABQ04091; NID:G9951872; PIDN:AAG0892
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA5536
C;Superfamily: Escherichia coli dosage-dependent dnaK suppressor protein dksA

Query Match 27.3%; Score 59; DB 2; Length 134;
Best Local Similarity 53.6%; Pred. No. 1.1;
Matches 15; Conservative 4; Mismatches 5; Indels 4; Gaps 2;

QY 6 BF-QLRDL---ADAWDLSSRSRQGWQMP 29

DB 41 EFGELDLRPSDEADLASREQROWL 68

RESULT 8
AH1130
glutamate decarboxylase homolog lmo0447 [imported] - Listeria monocytogenes (strain EGD
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AH1130
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, F
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1130
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-462 <GLA>
A;Cross-references: UNIPROT:O9F5P3; GB:NC_003210; PIDN:CAC98526.1; PID:G16409824; GSPDE
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0447
C;Superfamily: Escherichia coli glutamate decarboxylase

Query Match 26.4%; Score 57; DB 2; Length 462;
Best Local Similarity 42.9%; Pred. No. 7.1;
Matches 12; Conservative 8; Mismatches 4; Indels 4; Gaps 2;

QY 7 FQLRDLAD-AW---DLSSRSRQGWQMP 30
DB 374 YKLKDDANVAVTLYDLADLQMKGWQVP 401

RESULT 9
AC1370
glutamate decarboxylase homolog lmo2363 [imported] - Listeria monocytogenes (strain EGD
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC1370
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1370
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-464 <GLA>
A;Cross-references: UNIPROT:O9EYW9; GB:NC_003210; PIDN:CAD00441.1; PID:G16411851; GSPDE
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2363
C;Superfamily: Escherichia coli glutamate decarboxylase

Query Match 26.4%; Score 57; DB 2; Length 464;
Best Local Similarity 39.3%; Pred. No. 7.2;
Matches 11; Conservative 8; Mismatches 5; Indels 4; Gaps 1;

QY 7 FQLRDLADA----WDLSSRSRQGWQMP 30
DB 376 YKLKDDANVAVTLYDLADLQMKGWQVP 403

RESULT 10
AB1740
glutamate decarboxylase homolog lin2463 [imported] - Listeria innocua (strain Clip11262)

A;Molecule type: DNA
A;Residues: 1-739 <BEV>
A;Cross-references: UNIPROT:065440; EMBL:AL022603
A;Experimental source: cultivar Columbia; BAC clone F1935
A;Residues: 1-2034 <WIL>
A;Cross-references: UNIPROT:Q17388; EMBL:Z79755; PIDN:CAB02109.1; GSPDB:GN000019; CESP:F43G9
A;Experimental source: clone F43G9
C;Genetics:

```

A:Gene: CESP:F43G9.6
A:Map position: 1
A:Introns: 23/3; 65/3; 163/3; 258/3; 330/3; 400/2; 437/3; 578/2; 706/3; 919/3; 952/2; 11
Query Match      25.9%; Score 56; DB 2; Length 2034;
Best Local Similarity 44.0%; Pred. No. 42;
Matches 11; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 13 ADAMDLSRSRQRGWQMPNCRSRRG 37
   |||::|||::|||::|||
Db 1828 SDGVYSTKQRQGVKCPKCCTRRG 1852

RESULT 15
AE3623
Glutamate decarboxylase (EC 4.1.1.15) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AE3623
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
  Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
  Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3623
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-304 <KUP>
A:Cross-references: UNIPROT:Q8YBJ0; GB:AE008918; PIDN:AAL54152.1; PID:G17985117; GSPDB:G
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10910
A:Map position: II
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match      25.5%; Score 55; DB 2; Length 304;
Best Local Similarity 42.9%; Pred. No. 8.7;
Matches 12; Conservative 5; Mismatches 7; Indels 4; Gaps 1;

QY 7 FQRLDLAD---AWDLSSRSRQRGWQMP 30
   |||::|||::|||::|||
Db 217 FRIREGDPGYSLYDLSERLRLTGWQVP 244

```

Search completed: November 10, 2004, 14:52:28
 Job time : 9.56604 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: November 10, 2004, 13:38:57 : Search time 51.9509 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-62
Perfect score: 216
Sequence: 1 QRRGNEQLRLADAWDLSSRSRQRGWQMPNCRSRRGPG 39

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	64	29.6	466	1	DCEA_ECOLI	P80063 escherichia
2	64	29.6	466	1	DCEB_ECOLI	P28302 escherichia
3	64	29.6	466	2	Q7UAY1	Q7uay1 shigella fl
4	64	29.6	487	2	Q83PR1	Q83pr1 shigella fl
5	64	29.6	489	2	Q8FHG5	Q8fhg5 escherichia
6	60.5	28.0	234	2	O00833	O00833 trypanosoma
7	60.5	28.0	269	2	Q8CBH3	Q8cbh3 mus musculus
8	60	27.8	220	2	Q8C9I7	Q8c9i7 yarrowia li
9	60	27.8	466	1	DCEA_ECOS7	P82228 escherichia
10	59.5	27.5	730	2	Q8IVF8	Q8ivf8 homo sapien
11	59	27.3	130	2	Q6ZSH0	Q6zsh0 homo sapien
12	59	27.3	130	2	BAC86983	Bac86983 homo sapi
13	59	27.3	134	2	Q8HT38	Q8ht38 pseudomonas
14	58	26.9	168	2	Q8GNG6	Q8gng6 rhizobium l
15	57.5	26.6	82	2	Q853N5	Q853n5 mycobacteri
16	57.5	26.6	120	2	Q8T539	Q8t539 plasmodium
17	57	26.4	409	2	Q8C0D6	Q8c0d6 mus musculus
18	57	26.4	409	2	Q8C0F9	Q8c0f9 mus musculus
19	57	26.4	409	2	Q8COL5	Q8col5 mus musculus
20	57	26.4	462	1	DCEA_LISMO	Q9f5p3 listeria mo
21	57	26.4	464	1	DCEB_LISIN	Q928r9 listeria in
22	57	26.4	464	1	DCEB_LISMO	Q9eyw9 listeria mo
23	57	26.4	464	2	Q8GF15	Q8gf15 edwardsiell
24	57	26.4	464	2	Q7IX65	Q7ix65 listeria mo
25	57	26.4	464	2	AAT05100	Aat05100 listeria
26	57	26.4	479	2	Q7V891	Q7v891 prochloroco
27	57	26.4	560	2	Q7T0S1	Q7t0s1 xenopus lae
28	57	26.4	982	2	Q6MVK1	Q6mki bdellovibri
29	57	26.4	982	2	C8E79150	C8e79150 bdellovib
30	56.5	26.2	116	2	Q6ZF07	Q6zf07 oryza sativ
31	56.5	26.2	116	2	BAC83424	Bac83424 oryza sat

RESULT 1					
DCEA_ECOLI	STANDARD;				
ID_DCEA_ECOLI	STANDARD;	PRT;	466 AA.		
AC	P80063;				
DT	01-MAR-1992 (Rel. 21, Created)				
DT	01-DEC-1992 (Rel. 24, Last sequence update)				
DT	05-JUL-2004 (Rel. 44, Last annotation update)				
DE	Glutamate decarboxylase alpha (EC 4.1.1.15) (GAD-alpha).				
GN	Name=gadA; Synonyms=gadS; OrderedLocusNames=b3517, c4328;				
OS	Escherichia coli, and				
OS	Escherichia coli, O6.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Escherichia.				
OX	NCBI_TaxID=562, 217992;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=K12;				
RX	MEDLINE=92394884; PubMed=1522060;				
RA	Smith D.K., Kassam T., Singh B., Elliott J.F.;				
RT	"Escherichia coli has two homologous glutamate decarboxylase genes				
RT	that map to distinct loci.";				
RL	J. Bacteriol. 174:5820-5826(1992).				
RN	[2]				
RP	SEQUENCE, AND SEQUENCE OF 148-466 FROM N.A.				
RC	STRAIN=ATCC 11246;				
RX	MEDLINE=92155241; PubMed=1740158;				
RA	Maras B., Sweeney G., Barra D., Bossa F., John R.A.;				
RT	"The amino acid sequence of glutamate decarboxylase from Escherichia				
RT	coli. Evolutionary relationship between mammalian and bacterial				
RT	enzymes.";				
RL	Eur. J. Biochem. 204:93-98(1992).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=K12 / MG1655;				
RX	MEDLINE=94316500; PubMed=8041620;				
RA	Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;				
RT	"Analysis of the Escherichia coli genome. V. DNA sequence of the				
RT	region from 76.0 to 81.5 minutes.";				
RL	Nucleic Acids Res. 22:2576-2586(1994).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;				
RX	MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;				
RA	Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,				
RA	Rasko D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,				
RA	Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.F.,				
RA	Mobley H.L.T., Donnenberg M.S., Blattner F.R.;				
RT	"Extensive mosaic structure revealed by the complete genome sequence				
RT	of uropathogenic Escherichia coli.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).				
RN	[5]				
RP	SEQUENCE OF 1-22.				
RC	MEDLINE=94033862; PubMed=7764225;				
RA	Yoshida T., Yamashino T., Ueguchi C., Mizuno T.;				

Q9snf6 oryza sativ
Q7xik6 oryza sativ
Q8lnk1 oryza sativ
Q8h9f6 oryza sativ
Q74ki6 lactobacill
Aas08584 lactobaci
Q6c8t3 yarrowia li
Q6cr26 kluyveromyc
Q65420 arabidopsis
P55265 homo sapien
Q17388 caenorhabdi
Q9lg16 chilo iride
Q6nhv2 corynebacte
Cae49549 corynebac

RT "Expression of the *Escherichia coli* dimorphic glutamic acid
 RT decarboxylases is regulated by the nucleoid protein H-NS.";
 RL Biosci. Biotechnol. Biochem. 57:1568-1569(1993).
 RN [6]
 RP SEQUENCE OF 382-392.
 RC STRAIN=K12 / EMG2;
 RA MEDLINE=97443975; PubMed=9298646;
 RX Link A.J., Robinson K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of *Escherichia coli* K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 CC !- FUNCTION: Catalyzes the production of GABA.
 CC !- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
 CC !- COFACTOR: Pyridoxal phosphate.
 CC !- SUBUNIT: Homohexamer.
 CC !- SIMILARITY: Belongs to the group II decarboxylase family.
 CC !- DATABASE: NAME=Worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/GLDP/".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: M84024; AAA23833.1; -
 CC EMBL: X63123; CAA44834.1; -
 CC EMBL: U00039; AAB18493.1; -
 CC EMBL: A5000428; AAC76542.1; -
 CC EMBL: A5016768; AAN82764.1; ALT_INIT.
 CC F01: S47737; S24234.
 CC EC02DBASE; D046.5; 6TH EDITION.
 CC EC02DBASE; E046.5; 6TH EDITION.
 CC EC02DBASE; E04302; -
 CC EcoGene; EGS0009; Gada.
 CC InterPro; IPR010107; Glu decarb GAD.
 CC InterPro; IPR002129; Pyr-Idoxal dec.
 CC Pfam; PF00282; Pyridoxal dec; 1.
 CC TIGRPFAM; TIGR01788; Glu-decarb-GAD; 1.
 CC PROSITE; P500392; DDC GAD_HDC_YDC; 1.
 CC Complete proteome; Decarboxylase; Direct protein sequencing; Lyase;
 CC Multigene family; Pyridoxal phosphate.
 CC BINDING 276 276
 CC FT BINDING 64 64
 CC FT CONFLICT 73 73
 CC FT CONFLICT 153 153
 CC FT CONFLICT 165 165
 CC FT CONFLICT 208 208
 CC FT CONFLICT 295 295
 CC FT CONFLICT 355 355
 CC SEQUENCE 466 AA; 52685 MW; 86F963B710553E22 CRC64;
 Query Match 29.6%; Score 64; DB 1; Length 466;
 Best Local Similarity 50.0%; Pred. No. 4.7;
 Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;
 QY 7 FQRLDLAD---AWDLSSRSRGRQWMP 30
 Db 380 FLKQGEDPGYTLDSLRLRGKQVP 407
 RESULT 2
 DCBE_ECOLI STANDARD; PRT; 466 AA.
 AC P28302; P76873;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Glutamate decarboxylase beta (EC 4.1.1.15) (GAD-beta).
 GN Name=gadB; OrderedLocNames=b1493, 22215, EGS2098, SF1734, S1867;
 OS *Escherichia coli*,
 RT

OS *Escherichia coli* O157:H7, and
 OS *Shigella flexneri*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 83334, 623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12;
 RX MEDLINE=92394894; PubMed=1522060;
 RA Smith D.K., Kassam T., Singh B., Elliott J.F.;
 RT "Escherichia coli has two homologous glutamate decarboxylase genes
 RT that map to distinct loci.";
 RL J. Bacteriol. 174:5820-5826(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kunara S., Shiba T., Hattori M., Shingawa H.;
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RN [6]
 RP SEQUENCE OF 1-318 FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12;
 RA Turlin E., Gasser F., Biville F.;
 RT "Sequence and functional analysis of an *Escherichia coli* DNA fragment
 RT able to complement pqgE and pqgF from *Methylobacterium organophilum*.";
 RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1-15.
 RC SPECIES=E.coli; STRAIN=K12;
 RX MEDLINE=93204884; PubMed=8455549;
 RA Yoshida T., Ueguchi C., Yamada H., Mizuno T.;
 RT "Function of the *Escherichia coli* nucleoid protein, H-NS: molecular
 RT analysis of a subset of proteins whose expression is enhanced in a hns

RESULT 3	Q7UAY1	PRELIMINARY;	PRT;	466 AA.
AC	Q7UAY1			
DT	01-OCT-2003 (T-EMBLrel. 25, Created)			
DT	01-OCT-2003 (T-EMBLrel. 25, Last sequence update)			
DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)			
DE	Glutamate decarboxylase isozyme.			
GN	Name=gadA; OrderedLocusNames=S4173;			
OS	Shigella flexneri.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Shigella.			
OX	NCBI_TaxID=623;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=2457T;			
RX	MEDLINE=22590274; PubMed=12704152;			
RA	Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,			
RA	Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,			
RA	Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,			
RA	Schwartz D.C., Blattner F.R.;			
RT	"Complete genome sequence and comparative genomics of Shigella			
RT	flexneri serotype 2a strain 2457T.";			
RL	Infect. Immun. 71:2775-2786(2003).			
CC	-I- COFACTOR: Pyridoxal phosphate (By similarity).			
CC	-I- SIMILARITY: Belongs to the group II decarboxylase family.			
DR	EWEL; AE016391; AAP19142.1; --			
DR	GO; GO:0016831; F:carboxy-lyase activity; IEA.			
DR	GO; GO:0006520; P:amino acid metabolism; IEA.			
DR	InterPro; IPR010107; Glu decarb GAD.			
DR	InterPro; IPR001219; Pyridoxal dec.			
DR	InterPro; IPR001638; SBP bac 3			
DR	Pfam; PF00282; Pyridoxal_dec; 1			
DR	TIGRfams; TIGR01788; Glu-decarb-GAD; 1.			
DR	PROSITE; PS00392; DDC_GAD_HDC_YDC; UNKNOWN_1.			
DR	PROSITE; PS01039; SBP_BACTERIAL_3; UNKNOWN_1.			
KW	Lysase; Pyridoxal phosphate.			
SQ	SEQUENCE 466 AA; 52700 MW; 44B04876B2DF1680 CRC64;			
Query Match	29.6%;	Score 64;	DB 2;	Length 466;
Best Local Similarity	50.0%;	Pred. No. 4.7;		
Matches 14;	Conservative	4;	Mismatches 6;	Indels 4;
Gaps				
Qy	7 FOURDLAD----	AWDLSSRSRQGMQMP 30		
Db	380 FKLKDGEDPGYTLIDLSRLRLRGQVP 407			
RESULT 4				
Q83PR1	PRELIMINARY;	PRT;	487 AA.	
ID	Q83PR1			
AC	Q83PR1;			
DT	01-JUN-2003 (T-EMBLrel. 24, Created)			
DT	01-JUN-2003 (T-EMBLrel. 24, Last sequence update)			
DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)			
DE	Glutamate decarboxylase isozyme.			
GN	Name=gadA; OrderedLocusNames=SF3594;			
OS	Shigella flexneri.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Shigella.			
OX	NCBI_TaxID=623;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=301 / Serotype 2a;			
RX	MEDLINE=22272406; PubMed=12384590;			
RA	Jin O., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,			
RA	Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,			
RA	Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding X., Chen S.,			
RA	Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,			
RA	Xu J.;			
RT	"Genome sequence of Shigella flexneri 2a: insights into pathogenicity			
RT	through comparison with genomes of Escherichia coli K12 and O157.";			

```

RL Nucleic Acids Res. 30:4432-4441(2002).
CC -I- COFACTOR: Pyridoxal phosphate (By similarity).
CC -I- SIMILARITY: Belongs to the group II decarboxylase family.
DR EMBL; AE015367; AAN5045.1; -.
DR GO; GO:0016831; P:carboxy-lyase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR010107; Glu decarb GAD.
DR InterPro; IPR002129; Pyridoxal dec.
DR InterPro; IPR001638; SBP bac 3.
DR Pfam; PF00282; Pyridoxal_dec_1.
DR TIGRFAMs; TIGR01788; Glu-decarb-GAD; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; UNKNOWN_1.
DR PROSITE; PS01039; SBP_BACTERIAL_3; UNKNOWN_1.
KW Complete proteome; Lyase; Pyridoxal phosphate.
SQ SEQUENCE 487 AA; 55217 MW; 1B9D98041A7C38C2 CRC64;

Query Match 29.6%; Score 64; DB 2; Length 487;
Best Local Similarity 50.0%; Pred. No. 5;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLAD----AWDLSSRSRQGWMP 30
DB 401 FKLKGDGDPGYTLDSLRLRLRGWQP 428

RESULT 5
QBPHGS PRELIMINARY; PRT; 489 AA.
AC QBPHGS;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Glutamate decarboxylase beta (EC 4.1.1.15).
GN Name=gadB; OrderedLocustNames=cl922;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA McInley H.L.F., Domeneberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -I- COFACTOR: Pyridoxal phosphate (By similarity).
CC -I- SIMILARITY: Belongs to the group II decarboxylase family.
DR EMBL; AE016760; AAN80380.1; -.
DR GO; GO:0004353; P:glutamate decarboxylase activity; IEA.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR010107; Glu decarb GAD.
DR InterPro; IPR002129; Pyridoxal dec.
DR InterPro; IPR001638; SBP bac 3.
DR Pfam; PF00282; Pyridoxal_dec_1.
DR TIGRFAMs; TIGR01788; Glu-decarb-GAD; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; UNKNOWN_1.
DR PROSITE; PS01039; SBP_BACTERIAL_3; UNKNOWN_1.
KW Complete proteome; Lyase; Pyridoxal phosphate.
SQ SEQUENCE 489 AA; 55440 MW; D885B9796EBCAF7 CRC64;

Query Match 29.6%; Score 64; DB 2; Length 489;
Best Local Similarity 50.0%; Pred. No. 5;
Matches 14; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 7 FQLRDLAD----AWDLSSRSRQGWMP 30
DB 403 FKLKGDGDPGYTLDSLRLRLRGWQP 430

```

```

RESULT 6
QO0833 PRELIMINARY; PRT; 234 AA.
ID QO0833;
AC QO0833;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Trypanosoma cruzi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CL Brener;
RX MEDLINE=20160916; PubMed=10688909;
RA Vazquez M., Ben-Dov C., Lorenzi H., Moore T., Schijman A.G.,
RA Levin M.J.;
RT "The short interspersed repetitive element of Trypanosoma cruzi, SIRE,
RT is part of VIPER, an unusual retroelement related to long terminal
RT repeat retrotransposons.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2128-2133(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CL Brener;
RA Vazquez M.P.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12774; CAA73302.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 234 AA; 25905 MW; B900F6BB42652916 CRC64;

Query Match 28.0%; Score 60.5; DB 2; Length 234;
Best Local Similarity 44.8%; Pred. No. 6.7;
Matches 13; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

QY 3 RGNFQLDLADWDL-SSRSRQGWMP 30
DB 201 RGRVFTQLDLAKGWDLRGAGSGCWRT 229

RESULT 7
QBCHB3 PRELIMINARY; PRT; 269 AA.
ID QBCHB3;
AC QBCHB3;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:9630027A13 product:RABSA, member RAS oncogene
DE family, full insert sequence. (Fragment).
GN Name=Rab5a;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

```

RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20530913; PubMed=11079861;
 RA Shibata K., Itoh M., Azawa K., Nagaoaka S., Saeki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RA Adachi J., Azawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Satoh K., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK036011; BAC29275.1; -
 DR MGD; MGI:105926; Rab5a.
 DR GO; GO:0005763; C:early endosome; IDA.
 DR GO; GO:0030139; C:cytosolic vesicle; IDA.
 DR GO; GO:0045121; C:lipid raft; IDA.
 DR GO; GO:0001726; C:ruffles; IDA.
 DR GO; GO:0003924; F:GTPase activity; IDA.
 DR GO; GO:0019001; F:guanylate nucleotide binding; IDA.
 DR GO; GO:0006897; P:apoptosis; IDA.
 FT NON TER 1
 SQ SEQUENCE 269 AA; 2676 MW; 1FF74B51B36340CD CRC64;
 Query Match 28.0%; Score 60.5; DB 2; Length 269;
 Best Local Similarity 50.0%; Pred. No. 7.8;
 Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
 QY 15 AWDLSRSRQGWQ-MNCRSRRG 37
 Db 144 AWDGNPLRGRGWPLPFTWRRG 167
 RESULT 8
 Q6C917 PRELIMINARY; PRT; 220 AA.
 ID Q6C917
 AC Q6C917;
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Similar to sp|P21374|Saccharomyces cerevisiae YJR050w UTR3 pre-mRNA
 DE splicing factor.

GN ORFNames=YALI0D109019;
 OS Yarrowia lipolytica (Candida lipolytica).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxID=4952;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG GENOEVUES;
 RA Dujon B., Sherman D., Fischer G., Durkens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boistrame A., Boyer J., Cattolico L., Confanioli F., de Darivar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Karrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicolaud J.M., Nikolski M., Oxtas S., Ozier-Kalogeropoulos C.,
 RA Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swenne D., Toka F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudon B., Scarpelli C., Gaillardin C., Weissenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts.";
 RL Nature 430:35-44 (2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RA Genoscope;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR382130; CAG80863.1; -; B172423C76C30B7B CRC64;
 SQ SEQUENCE 220 AA; 25196 MW; B172423C76C30B7B CRC64;
 Query Match 27.8%; Score 60; DB 2; Length 220;
 Best Local Similarity 33.9%; Pred. No. 7.3;
 Matches 19; Conservative 6; Mismatches 9; Indels 22; Gaps 3;
 QY 5 NEFQLDLAD-----AWDLSSRS-----RQGWQNPNCRSRGPG 39
 Db 70 SDFQLRLDNLKMSERHRLDQLIRSLGGPNYASFGGKRGYQYVG-RARELPG 124
 RESULT 9
 DCEA_ECO57 STANDARD; PRT; 466 AA.
 ID DCEA_ECO57
 AC P58228;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Glutamate decarboxylase alpha (EC 4.1.1.15) (GAD-alpha).
 GN Name-gada; Synonyms=gads; OrderedLocusNames=z4930, EC44397;
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouisis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,

RA Iida T., Takami H., Honda T., Sasaoka C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,
 RT "Complete genome sequence of enterohemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 CC -!- FUNCTION: Catalyzes the production of GABA (By similarity).
 CC -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
 CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -!- SUBUNIT: Homohexamer (By similarity).
 CC -!- SIMILARITY: Belongs to the group II decarboxylase family.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB005577; AAG58658.1; -;
 CC EMBL; AF002565; BAB37820.1; -;
 CC PIR; E91178; E91178.
 CC PIR; F86024; F86024.
 CC InterPro; IPR010107; Gly decarb GAD.
 CC InterPro; IPR002129; Pyridoxal dec.
 CC Pfam; PF00282; Pyridoxal dec; 1.
 CC TIGRFAMs; TIGR01788; Glu-decarb-GAD; 1.
 CC PROSITE; P500392; DDC GAD_HDC YDC; 1.
 KW Complete proteome; Decarboxylase; Lyase; Multigene family;
 FT Pyridoxal phosphate.
 FT BINDING 276 Pyridoxal phosphate (By similarity).
 SQ SEQUENCE 466 AA; 52699 MW; C7F9623DBE24E489 CRC64;

Query Match 27.8%; Score 60; DB 1; Length 466;
 Best Local Similarity 46.4%; Pred. No. 16;
 Matches 13; Conservative 5; Mismatches 6; Indels 4; Gaps 1;
 QY 7 FQRLRLAD----AWDLSSRSRQGWMP 30
 DB 380 FKLKEGEDPGYTLYDLSEKRLRUGWQVP 407

RESULT 10
 Q81VF8 PRELIMINARY; PRT; 730 AA.
 AC Q81VF8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE KIA2013 protein (Fragment).
 GN Name-KIA2013.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Nagase T., Kikuno R., Ohara O.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB095933; BAC23109.1; -;
 FT NON_TER 1
 SQ SEQUENCE 730 AA; 79351 MW; 0D048672E371AEC8 CRC64;

Query Match 27.5%; Score 59.5; DB 2; Length 730;
 Best Local Similarity 46.2%; Pred. No. 31;
 Matches 12; Conservative 4; Mismatches 9; Indels 1; Gaps 1;
 QY 14 DAWDLSSRSRQGWMPNCRSRGPG 39
 DB 3 DGWPPRRRRKQGWRLLR-LLRGPG 27

RESULT 11
 Q6ZSHO PRELIMINARY; PRT; 130 AA.
 AC Q6ZSHO;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ45539.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thalamus;
 RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Masuho Y., Nagai K., Isogai T.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK127447; BAC86983.1; -;
 SQ SEQUENCE 130 AA; 14426 MW; 4FE3D779B71989AF CRC64;

Query Match 27.3%; Score 59; DB 2; Length 130;
 Best Local Similarity 32.1%; Pred. No. 5.7; 14; Indels 0; Gaps 0;
 Matches 9; Conservative 5; Mismatches 14; Indels 0; Gaps 0;
 QY 5 NEFQLRLADAWDLSSRSRQGWMPNC 32
 DB 19 NSLEAPYLVSGWGWGSRERKGWELITC 46

RESULT 12
 BAC86983 PRELIMINARY; PRT; 130 AA.
 ID BAC86983
 AC BAC86983;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE CDNA FLJ45539 fis, clone BRTHA203312.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thalamus;
 RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Masuho Y., Nagai K., Isogai T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK127447; BAC86983.1; -;
 SQ SEQUENCE 130 AA; 14426 MW; 4FE3D779B71989AF CRC64;

Query Match 27.3%; Score 59; DB 2; Length 130;
 Best Local Similarity 32.1%; Pred. No. 5.7; 14; Indels 0; Gaps 0;
 Matches 9; Conservative 5; Mismatches 14; Indels 0; Gaps 0;
 QY 5 NEFQLRLADAWDLSSRSRQGWMPNC 32
 DB 19 NSLEAPYLVSGWGWGSRERKGWELITC 46

RESULT 13
 Q9HT38

```

RA Mochizuki Y., Nakayama S.; Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
DR EMEL; AP003001; BAB50180.1; -.
DR HSSP; P32939; 1KY3.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR InterPro; IPR001806; Ras.transfmng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; Ras.1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR TIGRFAMs; TIGR00231; small_GTP.1.
KW Complete proteome; GTP-binding.
SQ SEQUENCE 168 AA; 18516 MW; 47F9B229D3EF0850 CRC64;

Query Match 26.9%; Score 58; DB 2; Length 168;
Best Local Similarity 44.4%; Pred.No.10;
Matches 12; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

QY 7 FQRLDLADWLDS----SRSRQGWQM 29
DB 118 FNKSLADRWELSDGSDIDNLRQLGWQI 144
|||||:|:|
|||||:|:|

RESULT 15
Q853N5 PRELIMINARY; PRT; 82 AA.
AC Q853N5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp42.
OS Mycobacteriophage Bx1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
CX NCBI_TaxID=205877;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22592660; PubMed=12705866;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003)
DR EMBL; A129337; AN16702.1; -.
SQ SEQUENCE 82 AA; 9442 MW; 9D58148FEA7486EC CRC64;

Query Match 26.6%; Score 57.5; DB 2; Length 82;
Best Local Similarity 48.6%; Pred.No.5.6;
Matches 17; Conservative 1; Mismatches 12; Indels 5; Gaps 4;

QY 6 EFQLRDL-ADADWLSSRSRQGWQM--PNCRSRRG 37
DB 49 EAAAEELRDGW-LVLRER-RGWSLICPNCRKREG 81
|||||
|||||
|||||
|||||

Search completed: November 10, 2004, 14:50:34
Job time : 53.9509 secs

```

This Page Blank (uspio)